

## RESEARCH ARTICLE

# Targeted dimensionality reduction enables reliable estimation of neural population coding accuracy from trial-limited data

Charles R. Heller<sup>1,2\*</sup>, Stephen V. David<sup>2\*</sup>

**1** Neuroscience Graduate Program, Oregon Health and Science University, Portland, Oregon, United States of America, **2** Oregon Hearing Research Center, Oregon Health and Science University, Portland, Oregon, United States of America

✉ Current address: Max Planck Institute for Biological Cybernetics, Tübingen, Germany  
\* [davids@ohsu.edu](mailto:davids@ohsu.edu)



## OPEN ACCESS

**Citation:** Heller CR, David SV (2022) Targeted dimensionality reduction enables reliable estimation of neural population coding accuracy from trial-limited data. PLoS ONE 17(7): e0271136. <https://doi.org/10.1371/journal.pone.0271136>

**Editor:** David S Vicario, Rutgers The State University of New Jersey, UNITED STATES

**Received:** January 23, 2022

**Accepted:** June 23, 2022

**Published:** July 21, 2022

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0271136>

**Copyright:** © 2022 Heller, David. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All data and code used for producing the figures and analysis in the manuscript is available on GitHub at <https://github.com/crheller/dDR>. We have also used Zenodo to

## Abstract

Rapidly developing technology for large scale neural recordings has allowed researchers to measure the activity of hundreds to thousands of neurons at single cell resolution *in vivo*. Neural decoding analyses are a widely used tool used for investigating what information is represented in this complex, high-dimensional neural population activity. Most population decoding methods assume that correlated activity between neurons has been estimated accurately. In practice, this requires large amounts of data, both across observations and across neurons. Unfortunately, most experiments are fundamentally constrained by practical variables that limit the number of times the neural population can be observed under a single stimulus and/or behavior condition. Therefore, new analytical tools are required to study neural population coding while taking into account these limitations. Here, we present a simple and interpretable method for dimensionality reduction that allows neural decoding metrics to be calculated reliably, even when experimental trial numbers are limited. We illustrate the method using simulations and compare its performance to standard approaches for dimensionality reduction and decoding by applying it to single-unit electrophysiological data collected from auditory cortex.

## Introduction

Neural decoding analysis identifies components of neural activity that carry information about the external world (e.g. stimulus identity). This approach can offer important insights into how and where information is encoded in the brain. For example, classic work by Britten et al. demonstrated that the ability of single neurons in area MT to decode visual stimuli closely corresponds to animal's perceptual performance [1]. Thus, by using decoding the authors identified a possible neural substrate for detection of motion direction [1]. Yet, behavior does not depend solely on single neurons. In the years since this work, many theoretical frameworks have been proposed for how information might be pooled across individual neurons into a population code [2–8]. One clear theme that has emerged from this work is that stimulus

assign a DOI to the repository: [10.5281/zenodo.5788573](https://doi.org/10.5281/zenodo.5788573).

**Funding:** This work was supported by a National Science Foundation Graduate Research Fellowship (NSF GRFP, GVPRS0015A2) (CRH), the National Institute of Health (NIH, R01 DC0495) (SVD), Achievement Rewards for College Scientists (ARCS) Portland chapter (CRH), and by the Tartar Trust at Oregon Health and Science University (CRH). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

independent, correlated activity (*i.e.* noise correlations) between neurons may substantially impact information coding [2, 4–8]. This has now been confirmed *in vivo* using decoding analysis to measure the information content of large neural populations [9–11]. Therefore, covariability between neurons must be taken into account when measuring population coding accuracy.

Under most experimental conditions, estimates of pairwise correlation between neurons is unreliable due to insufficient sampling (*e.g.* too few stimulus repeats) [12]. In these situations, traditional decoding algorithms are likely to over-fit to noise in the neural data. This issue becomes even more apparent as the number of pairwise interactions that must be estimated increases, a situation that is becoming more common due to the recent growth in large-scale neurophysiology techniques [13]. In some cases, *e.g.* for chronic recording experiments and anesthetized preps, the number of trials can be increased to circumvent this issue. However, in behavioral experiments, where the number of trials is often fundamentally limited by variables such as animal performance, new analytical techniques for decoding are required.

Here, we present decoding-based dimensionality reduction (*dDR*), a simple and generalizable method for dimensionality reduction that significantly mitigates issues around estimating correlated variability in experiments with a relatively low ratio of observations to neurons. Our method takes advantage of recent observations that population covariability is often low-dimensional [14–17] to define a subspace where decoding analysis can be performed reliably while still preserving the dominant mode(s) of population covariability. The *dDR* method can be applied to data collected across many different stimulus and/or behavior conditions, making it a flexible tool for analyzing a wide range of experimental data.

We motivate the requirement for dimensionality reduction by illustrating how estimates of a popular information decoding metric,  $d'^2$  [4, 5], can be biased by small experimental sample sizes. Building on a simple two-neuron example, we demonstrate that low-dimensional structure in the covariability of simulated neural activity can be leveraged to reliably decode stimulus information, even when the number of neurons exceeds the number of experimental observations. Finally, we use a dataset collected from primary auditory cortex to highlight the advantages of using *dDR* for neural population decoding over standard principal component analysis.

## Materials and methods

### Surgical procedure

All procedures were performed in accordance with the Oregon Health and Science University Institutional Animal Care and Use Committee (IACUC) and conform to standards of the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC). The surgical approach was similar to that described previously [18]. Adult male ferrets were acquired from an animal supplier (Marshall Farms). Head-post implantation surgeries were then performed in order to permit head-fixation during neurophysiology recordings. Two stainless steel head-posts were fixed to the animal along the midline using bone cement (Palacos), which bonded to the skull and to stainless steel screws that were inserted into the skull. After a two-week recovery period, animals were habituated to a head-fixed posture and auditory stimulation. At this point, a small (0.5–1 mm) craniotomy was opened above primary auditory cortex (A1) for neurophysiological recordings.

### Neurophysiology

Recording procedures followed those described previously [19, 20]. Briefly, upon opening a craniotomy, 1–4 tungsten micro-electrodes (FHC, 1–5 M $\Omega$ ) were inserted to characterize the

tuning and response latency of the region of cortex. Sites were identified as A1 by characteristic short latency responses, frequency selectivity, and tonotopic gradients across multiple penetrations [21]. Subsequent penetrations were made with a 64-channel silicon electrode array [22]. Electrode contacts were spaced 20  $\mu\text{m}$  horizontally and 25  $\mu\text{m}$  vertically, collectively spanning 1.05 mm of cortex. Data were amplified (RHD 128-channel headstage, Intan Technologies), digitized at 30 KHz (Open Ephys [23]) and saved to disk for further analysis.

Spikes were sorted offline using Kilosort2 (<https://github.com/MouseLand/Kilosort2>). Spike sorting results were manually curated in phy (<https://github.com/cortex-lab/phy>). For all sorted and curated spike clusters, a contamination percentage was computed by measuring the cluster isolation in feature space. All sorted units with contamination percentage less than or equal to 5 percent were classified as single-unit activity. All other stable units that did not meet this isolation criterion were labeled as multi-unit activity. Both single and multi-units were included in all analyses.

### Acoustic stimuli

Digital acoustic signals were transformed to analog (National Instruments), amplified (Crown), and delivered through a free-field speaker (Manger) placed 80 cm from the animal's head and 30° contralateral to the hemisphere in which neural activity was recorded. Stimulation was controlled using custom MATLAB software (<https://bitbucket.org/lbhb/baphy>), and all experiments took place inside a custom double-walled sound-isolating chamber (Professional Model, Gretch-Ken).

Auditory stimuli consisted of narrowband white noise stimuli with  $\approx 0.3$  octave bandwidth. In total, we presented fifteen distinct, non-overlapping noise bursts spanning a 5 octave range. Each noise was presented alone (-Inf dB) condition, or with a pure tone embedded at its center frequency for a range of different signal to noise ratios (-10dB, -5dB, 0dB). Thus, each experiment consisted of 60 unique stimuli (4 SNR conditions X 15 center frequencies). Overall sound level was set to 60 dB SPL. Stimuli were 300ms in duration with 200ms ISI and each sound was repeated 50 times per experiment in a pseudo-random sequence.

### Bootstrapped estimates of decoding performance for different sample sizes

In Fig 6 panels d and g, we present the relative performance of *dDR* vs. *taPCA* and *stPCA* applied to real neural data across different sample sizes. Unlike our simulations, here we were restricted to a finite number of total trials ( $k = 50$ ). Therefore, we utilized the following bootstrapping procedure to compute unbiased estimates of the standard error in cross-validated decoding performance at each sample size.

First, we selected a subset of the available trials ( $k = 15$ ) to hold out for validation. Next, we re-sampled, with replacement, from the remaining data to build bootstrapped estimation sets. For example, for  $k = 20$  this means that for each bootstrap sample we randomly selected 5 trials, with replacement, excluding the validation data. We then performed dimensionality reduction using *dDR*, *taPCA*, or *stPCA* and fit a decoding axis in the reduced dimensionality space for these 5 trials. Finally, we evaluated the decoding performance using this decoding axis on the held out  $k = 25$  validation trials. We normalized the resulting decoding metric,  $d'^2$ , for *taPCA* and *stPCA* to the mean *dDR*  $d'^2$  across all bootstraps for a given sample size. Thus, for each sample size  $k$ , we obtained a bootstrapped distribution of relative decoding performance between *dDR* and either *taPCA* or *stPCA*. The lines in Fig 6d, g represent the mean of this metric across bootstraps and the shading represents the standard deviation across bootstraps *i.e.*, the bootstrapped estimate of standard error [24]. For  $k < 50$  we performed a bootstrap

correction on the standard error in order to account for re-sampling only a subset of our full  $k = 50$  trials sample [25].

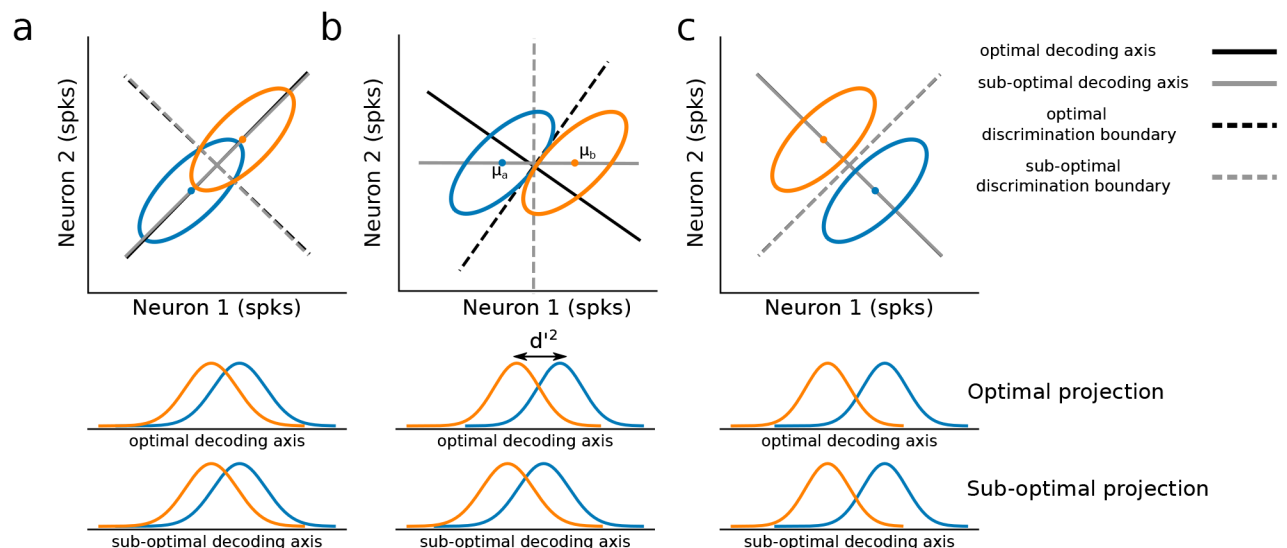
## Results

### Neural population decoding and noise correlations

Decades of neurophysiology experiments have demonstrated that neural activity under most experimental conditions is variable. For example, in the auditory system, the number of action potentials a neuron produces in response to a sound stimulus varies each time that sound is presented. This stimulus-independent variability has often been attributed to stochastic noise. However, it is increasingly appreciated that latent physiological processes, such as changes in arousal and attention, could be driving these apparently spontaneous fluctuations in neural responses [15, 17, 26, 27].

Supporting the hypothesis that latent processes drive variability, modern neural recording techniques have demonstrated that stimulus-independent variability is often correlated between neurons. Thus, it cannot simply be attributed to independent noise in each neuron. Because this variability is not related to an experimentally controlled variable, like stimulus condition, it is commonly referred to as noise correlation. A rich literature exists describing the importance of noise correlation for understanding neural population codes (for a review, see [7]). For the purposes of our work, it is useful to briefly highlight some of the main concepts and define key terminology that we use throughout this manuscript. Readers familiar with previous work on pairwise noise correlation may wish to skip to the next section.

Noise correlation can be visualized by plotting the response distribution of a pair of neurons in state-space (Fig 1a–1c), where state-space refers to the euclidean space in which each axis represents the activity of a single neuron. If two neurons share a noise correlation, as is shown in Fig 1a–1c, their response distribution (illustrated by ellipses) will be elongated. This



**Fig 1. Neuronal noise correlation and population coding.** a–c. In each panel, the spiking responses of two neurons are simulated for two experimental conditions (blue vs. orange). Noise correlation strength and the absolute distance between the mean responses were fixed across all simulations. Top: Response distributions in state-space under each condition are summarized using an ellipse that shows one standard deviation of responses around the mean. Middle: Responses are projected onto the optimal linear decoding axis, where decoding metrics such as  $d^2$  can be visualized and measured.  $d^2$  quantifies how discriminable two Gaussian distributions are. Bottom: Responses are projected onto a sub-optimal decoding axis. Unlike the optimal axis, this does not take into account noise correlations. Figure is adapted from Averbeck & Lee, 2006 [6].

<https://doi.org/10.1371/journal.pone.0271136.g001>

reflects the observation that when one neuron fires more spikes than average, the other neuron tends to do so as well. The sign and strength of a noise correlation is reflected by the shape of this response distribution.

Theoretical work has shown that the ability of a population of neurons to discriminate between different stimulus conditions critically depends on how noise correlation interacts with sensory tuning. To illustrate this point, Fig 1a–1c shows simulated response distributions for two neurons under two different stimulus conditions (blue vs. orange). When the noise correlation is aligned with the coding axis, it interferes with the ability to discriminate between the two distributions (Fig 1a). However, when the noise correlation is orthogonal, discrimination is actually easier (Fig 1c). This makes sense intuitively—if the uncontrolled variability (noise correlation) changes neural activity in the same way as the stimulus, then it is impossible to know if a change in the activity should be attributed to stimulus, or to noise.

In practice, noise correlation need not be perfectly aligned with, or orthogonal to, the coding axis. The specific alignment of noise correlation can be leveraged to achieve an optimal decoding strategy. The intuition for this optimization is illustrated in Fig 1b. The linear decoding axis is rotated to minimize the amount of noise correlation observed by a *e.g.* downstream readout neuron (Fig 1c, middle) relative to a sub-optimal decoding strategy that only takes into account the trial-averaged activity of each response distribution (Fig 1b, bottom). Thus, accurately measuring noise correlation is important for optimally decoding neural population activity.

### Small sample sizes limit the reliability of neural decoding analysis

Linear decoding identifies a linear, weighted combination of neural activity along which distinct experimental conditions (*e.g.* different sensory stimuli) can be discriminated. In neural state-space, this weighted combination is referred to as the decoding axis,  $w_{opt}$ , the line along which the distance between stimulus classes is maximized and trial-trial variance is minimized (Fig 2a and 2b). To quantify decoding accuracy, single-trial neural activity is projected onto this axis and a decoding metric is calculated to quantify the discriminability of the two stimulus classes. Here, we use  $d'^2$ , the discrete analog of Fisher Information [4, 5]. This discriminability metric has been used in a number of previous studies [6, 9–11, 28] and has a direct relationship to classical signal detection theory [4, 29].

Looking at the simulated data in Fig 2a and 2b, one can appreciate that an accurate estimate of  $w_{opt}$  requires knowledge of both the mean response evoked by each stimulus class ( $\mu_a$  vs.  $\mu_b$ ), as well as the population noise correlations,  $\Sigma$  (summarized by the ellipses in Fig 2a and 2b). Indeed,  $d'^2$ , is directly dependent on these features:

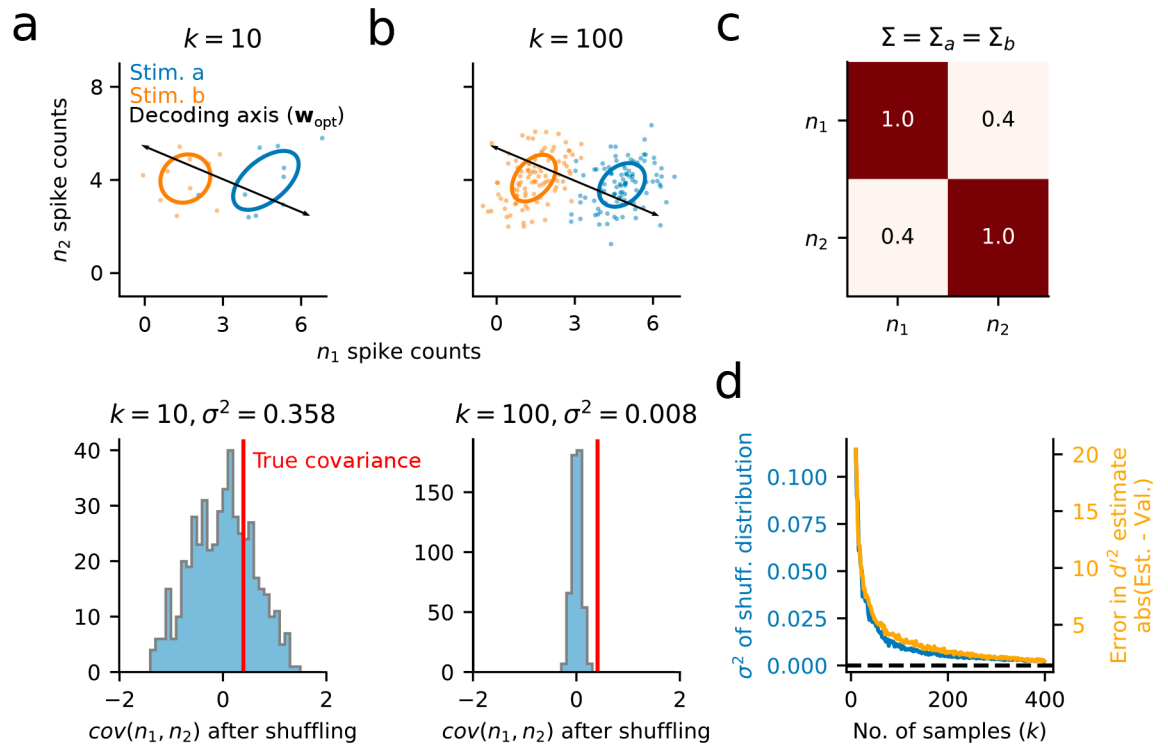
$$d'^2 = \Delta\mu^T w_{opt} \quad (1)$$

$$w_{opt} = \Sigma^{-1} \Delta\mu \quad (2)$$

$$\Delta\mu = \mu_a - \mu_b \quad (3)$$

Where  $\mu_a$  and  $\mu_b$  are the  $N \times 1$  vectors describing the mean response of an  $N$ -neuron population to two stimuli,  $a$  vs.  $b$ , respectively, and  $\Sigma$  is the average  $N \times N$  covariance matrix  $\frac{1}{2}(\Sigma_a + \Sigma_b)$  (*e.g.* Fig 2c).

In practice, the noise correlations between neurons (or  $r_{sc}$ ) is reported to be very small—on the order of  $10^{-1}$  or  $10^{-2}$  [30–32]. As we can see from the shuffled distribution in Fig 2a (bottom), this can pose a problem for accurate estimates of the off-diagonal elements in  $\Sigma$ , and, as a consequence,  $w_{opt}$  itself. This difficulty is especially pronounced when sample sizes are



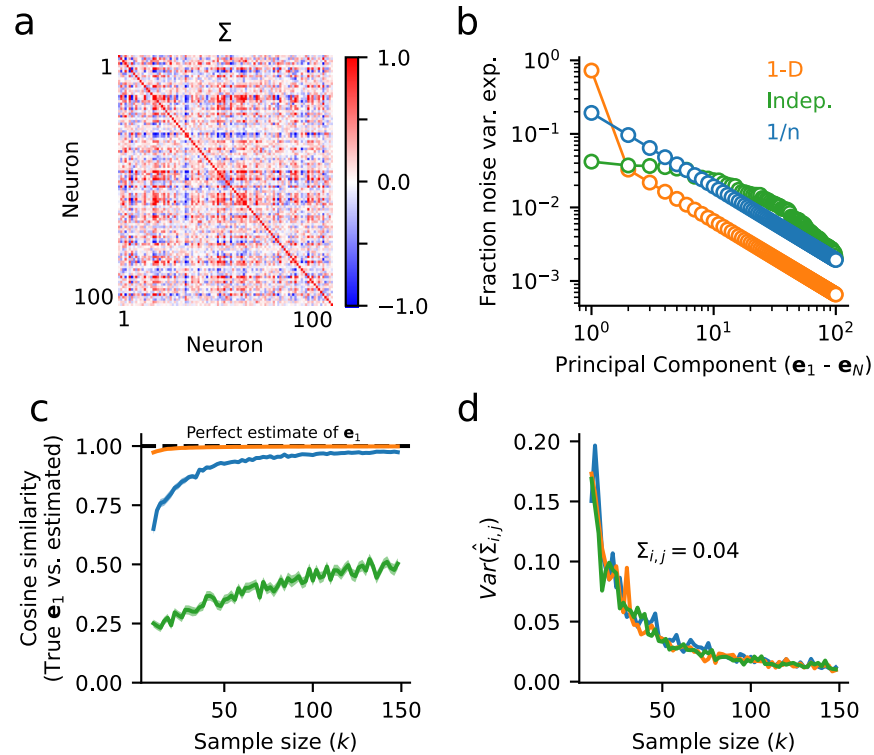
**Fig 2. Measurements of noise correlations and discriminability are unreliable when sampling is limited.** **a.** Top:  $k = 10$  single trial spike count responses are drawn from standard multivariate Gaussians  $\mathcal{N}(\mu_a, \Sigma)$  and  $\mathcal{N}(\mu_b, \Sigma)$  corresponding to two different stimulus conditions,  $a$  and  $b$ . Ellipses show the standard deviation of spike counts across trials. Bottom: Reliability of the noise correlation estimate between neuron 1 ( $n_1$ ) and neuron 2 ( $n_2$ ) is calculated by shuffling values of  $n_i$  500 times. The true covariance (red line) falls within this distribution, indicating that estimates of covariance are not reliable for  $k = 10$ . **b.** Same as in (a), but drawing  $k = 100$  samples for each stimulus. The narrower distribution of permuted measures indicates a greater likelihood of identifying an accurate estimate of covariance. **c.** The covariance matrix,  $\Sigma$ , used to generate data in (a)/(b). The true pairwise covariance for this pair of simulated neurons has a value of 0.4. **d.** Variance ( $\sigma^2$ ) of covariance estimates based on the permutation analysis in (a)/(b) for a range of sample sizes,  $k$  (blue). Variance decays as  $\mathcal{O}(\frac{1}{k-1})$  (see S1 Appendix). Overlaid is the difference in stimulus discriminability,  $d^2$  (Eq 1), between estimation and validation sets (50–50 split) estimated for each sample size (orange). Large values in the  $d^2$  difference for low  $k$  indicate overfitting of  $w_{opt}$  to the estimation data. This difference asymptotes toward zero as sample size increases and the estimate of covariance becomes reliable.

<https://doi.org/10.1371/journal.pone.0271136.g002>

relatively small (compare Fig 2a to 2b). The estimates of covariance and stimulus discriminability improve with increasing sample size, but robust performance is not reached until  $\approx 100$  stimulus repetitions, even for this case with relatively strong covariance (Fig 2d). The sample sizes (e.g. number of trials) in most experiments, especially those involving animal behavior, are typically much lower, raising the question: How can one reliably quantify coding accuracy in large neural populations observed over relatively few trials?

### Neural population activity is low-dimensional

Analysis of neural population data with dimensionality reduction has consistently revealed low-dimensional structure in neural activity [33]. Specifically, recent studies have found that this noise correlation is dominated by a small number of latent dimensions [14, 15, 17, 27]. Noise correlation impacts stimulus coding accuracy [7] and is known to depend on internal states, such as attention, that affect behavioral task performance [15, 16, 30, 34, 35]. These findings suggest that the space of neural activity relevant for understanding stimulus decoding,



**Fig 3. Low-dimensional noise correlation can be estimated reliably for neural populations, even when pairwise noise correlation cannot.** **a.** Example covariance matrix,  $\Sigma$ , for a 100-neuron population with low-dimensional covariance structure. **b.** Scree plot shows the fraction of total population variance captured along each noise dimension, computed by PCA, for three different datasets with varying dimensionality. Orange: 1-dimensional noise (1-D), covariance matrix in (a); green: independent noise (Indep.); blue: power law decay ( $1/n$ ). **c.** Surrogate datasets with varying numbers of samples,  $k$ , are drawn from the three noise distributions in (b). For each dataset, the cosine similarity between the estimate of the largest noise dimension,  $\hat{e}_1$ , and the true noise dimension,  $e_1$ , is plotted as function of sample size. For low-dimensional data,  $e_1$  can be estimated very reliably. **d.** Variance in the estimate of covariance,  $\hat{\Sigma}_{i,j}$ , for two neurons with a true covariance of 0.04 is plotted as a function of the number of trials, as in Fig 2d. Even at sample sizes  $>100$ ,  $Var(\hat{\Sigma}_{i,j}) \approx 0.02$ , corresponding to a standard deviation of  $\approx 0.14$ . Therefore, estimates of  $\hat{\Sigma}_{i,j}$  may be off by up to an order of magnitude. Note that the amount of uncertainty does not depend on the dimensionality of the data, and results for all three datasets overlap (see S1 Appendix for an analytical derivation).

<https://doi.org/10.1371/journal.pone.0271136.g003>

and its relationship to behavior, may be small relative to the total number of recorded neurons.

When population data exhibits low-dimensional structure, the largest eigenvectors of  $\Sigma$  (i.e. the top principal components of population activity) provide a reasonable, low-rank approximation to the full-rank covariance matrix. Importantly, these high variance dimensions of covariability can be estimated accurately even from limited samples. To illustrate this point, we simulated population spike counts,  $X$ , for  $N = 100$  neurons by drawing  $k$  samples from a multivariate Gaussian distribution with mean  $\mu$  and covariance  $\Sigma$  (Eq 4).

$$X = \mathcal{N}(\mu, \Sigma) + \epsilon_{indep}. \tag{4}$$

Where in Eq 4,  $\epsilon_{indep}$  represents a small amount of independent noise added to each neuron, effectively removing any significant structure in the smaller noise modes.

To investigate how different noise structures impact estimates of  $\Sigma$ , we simulated three different surrogate populations. First, we simulated data with just one large, significant noise dimension (Fig 3b–3d, 1-D data, orange). In this case, the first eigenvector can be estimated

reliably, even from just a few samples (Fig 3c). However, when the noise is independent and shared approximately equally across all neurons, estimates of the first eigenvector are poor (Fig 3c, Indep. noise, green). These first two simulations represent extreme examples—in practice, population covariability tends to be spread across at least a few significant dimensions [36]. To investigate a scenario that more closely mirrors this structure, we simulated a third dataset where the noise eigenspectrum decayed as  $1/n$ , where  $n$  goes from  $n = 1$  to  $N$ . Recent studies of large neural populations suggest that this power law relationship is a reasonable approximation to real neural data [36]. In this case, by  $k \approx 50$  trials, estimates of the first eigenvector are highly reliable, approaching a cosine similarity of  $\approx 0.9$  between the estimated and true eigenvectors (Fig 3c,  $1/n$  noise, blue). In all simulations, regardless of dimensionality, we find that estimates of single elements of  $\Sigma$  (i.e. single noise correlation coefficients) are highly unreliable (Fig 3d), as we see in the two-neuron example (Fig 2d).

Collectively, these simulations demonstrate that accurate estimates of noise correlation need not necessarily be limited by uncertainty in estimates of individual noise correlation coefficients themselves. In the following sections we describe a simple decoding-based dimensionality reduction algorithm, *dDR*, that leverages low-dimensional structure in neural population activity to facilitate reliable measurements of neural decoding.

### Decoding-based Dimensionality Reduction (*dDR*)

The *dDR* algorithm operates on a pairwise basis. That is, given a set of neural data collected over  $S$  different conditions, a different *dDR* projection exists for each of the  $\frac{S!}{2!(S-2)!}$  unique pairs. For simplicity, we will describe the case where  $S = 2$ , and consider these to be two unique stimulus conditions. However, note that the method can be applied in exactly the same manner to handle datasets with many different types and numbers of decoding conditions, where a unique *dDR* projection would then exist for each pair.

Let us consider the spiking response of an  $N$ -neuron population evoked by two different stimuli,  $S_a$  and  $S_b$ , over  $k$ -repetitions of each stimulus. From this data we form two response matrices,  $A$  and  $B$ , each with shape  $N \times k$ . Remembering that our goal is to estimate discriminability ( $d'^2$ , Eq 1), the *dDR* projection should seek to preserve information about both the mean response evoked by each stimulus condition,  $\mu_a$  and  $\mu_b$ , as well as the noise correlations,  $\Sigma$ . Therefore, we define the first dimension of *dDR* to be the axis that maximally separates  $\mu_a$  and  $\mu_b$ . We call this the *signal* axis.

$$\text{signal} = \mu_a - \mu_b = \Delta\mu \quad (5)$$

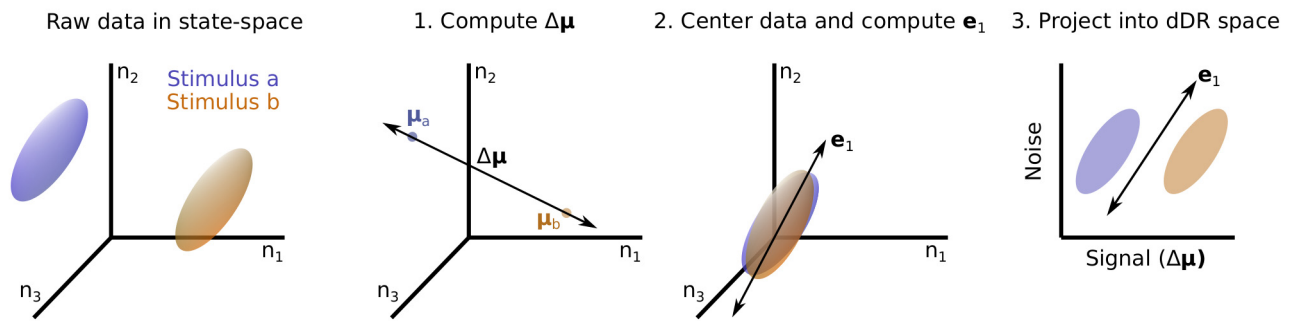
Next, we compute the first eigenvector of  $\Sigma$ ,  $e_1$ . This represents the largest noise mode of the neural population activity. Together, *signal* ( $\Delta\mu$ ) and  $e_1$  span the plane in state-space that is most optimized for reliable decoding. Finally, to form an orthonormal basis, we define the second *dDR* dimension as the axis orthogonal to  $\Delta\mu$  in this plane. As this second dimension is designed to preserve noise covariance, we call this the *noise*<sub>1</sub> axis.

$$\text{noise}_1 = e_1 - e_1\Delta\mu^T \quad (6)$$

The process outlined above is schematized graphically in Fig 4.

Thus, the *signal* and *noise*<sub>1</sub> axes make up a  $2 \times N$  set of weights, analogous to the loading vectors in standard PCA, for example. By projecting our  $N \times k$  data onto this new basis, we capture both the stimulus coding dimension ( $\Delta\mu$ ) and preserve the principal noise correlation dimension ( $e_1$ ), two critical features for measuring stimulus discriminability. Importantly, because  $e_1$  can be measured more robustly than  $\Sigma$  itself (Fig 3), performing this dimensionality





**Fig 4. Decoding-based Dimensionality Reduction (*dDR*).** Left to right: Responses of 3 neurons ( $n_1, n_2, n_3$ ) to two different stimuli are schematized in state-space. Ellipsoids illustrate the variability of responses across trials. **1.** To perform *dDR*, first the difference is computed between the two mean stimulus responses,  $\Delta\mu$ . **2.** Next, the mean response is subtracted for each stimulus to center the data around 0, and *PCA* is used to identify the first eigenvector of the noise covariance matrix,  $e_1$  (additional noise dimensions  $e_m, m > 1$  can be computed, see text). **3.** Finally, the raw data are projected onto the plane defined by  $\Delta\mu$  and  $e_1$ .

<https://doi.org/10.1371/journal.pone.0271136.g004>

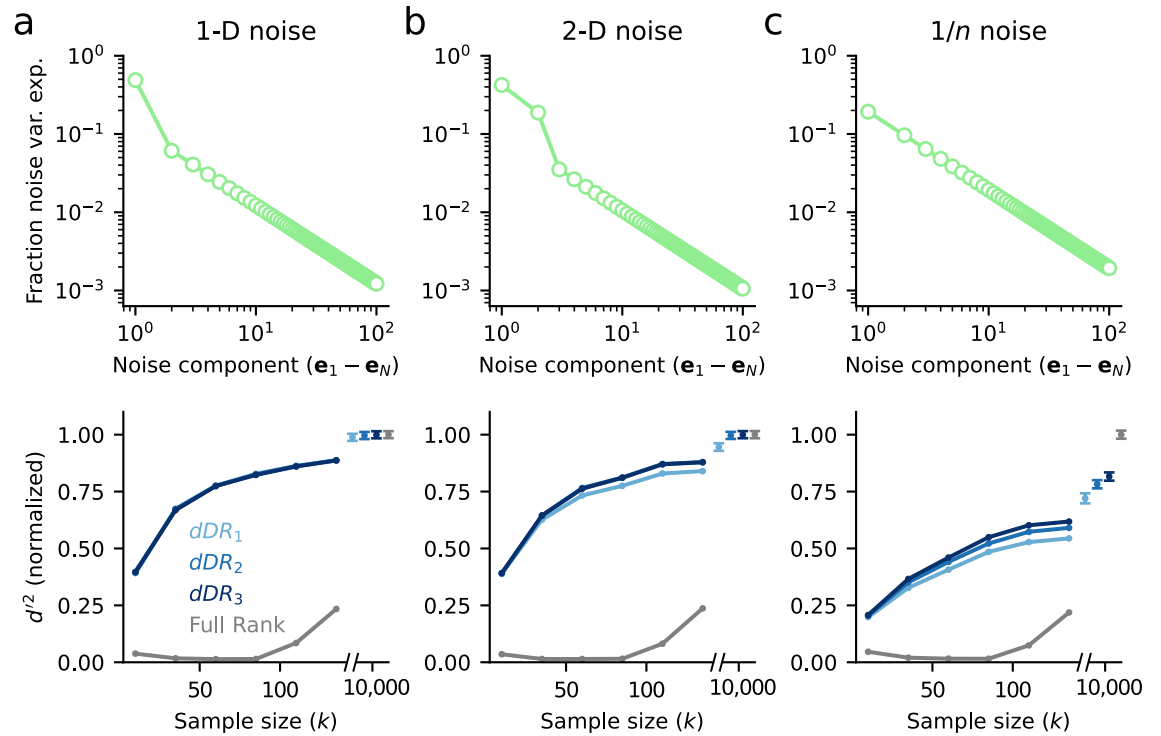
reduction helps mitigate the issues we encounter due to small sample sizes and large neural datasets.

As mentioned in the previous section, neural data often contains more than one significant dimension of noise correlation. To account for this, *dDR* can easily be extended to include more noise dimensions. To include additional dimensions, we deflate the spike count matrix,  $X$ , by subtracting out the *signal* and *noise*<sub>1</sub> dimensions identified by standard *dDR*, then perform *PCA* on the residual matrix to identify  $m$  further *noise* dimensions. Note, however, that for increasing  $m$  the variance captured by each dimension gets progressively smaller. Therefore, estimation of these subsequent noise dimensions becomes less reliable and will eventually become prone to over-fitting, especially with small sample sizes. For this reason, care should be taken when extending *dDR* in this way.

To demonstrate the performance of the *dDR* method, we generated three sample datasets containing  $N = 100$  neurons and  $S = 2$  stimulus conditions. Each of the three datasets contained unique noise correlation structure: 1.  $\Sigma$  contained one significant dimension (Fig 5a) 2.  $\Sigma$  contained two significant dimensions (Fig 5b) 3. Noise correlation decayed as  $1/n$  (Fig 5c). For each dataset, we measured cross-validated  $d^2$  between stimulus condition  $a$  and stimulus condition  $b$  using standard *dDR* with one noise dimension (*dDR*<sub>1</sub>), with two noise dimensions (*dDR*<sub>2</sub>), or with three noise dimensions (*dDR*<sub>3</sub>). We also estimated  $d^2$  using the full-rank data, without performing *dDR*. The bottom panels of Fig 5a–5c plot the decoding performance of each method as a function of sample size (*i.e.* number of stimulus repetitions). In each case,  $d^2$  is normalized to the asymptotic performance of the full-rank approach, when the number of samples is  $\gg$  than the number of neurons. This provides an approximate estimate of true discriminability for the population.

In contrast to the full-rank data where overfitting leads to dramatic underestimation of  $d^2$  on the test data for most sample sizes (Fig 5a–5c, bottom, grey lines), we find that  $d^2$  estimates after performing *dDR* are substantially more accurate and, critically, more reliable across sample sizes. That is, asymptotic performance of the *dDR* method is reached much more quickly than for the full-rank method.

For the one-dimensional noise case, note that there is no benefit of including additional *dDR* dimensions (Fig 5a), while for the higher dimensional data shown in Fig 5b and 5c, we see some improvements with *dDR*<sub>2</sub> and *dDR*<sub>3</sub>. However, these benefits don't begin to appear



**Fig 5. Evaluation of decoding accuracy and reliability with *dDR*.** **a.** Analysis of data with one-dimensional (1-D) noise covariance. For each sample size,  $k$ , 100 datasets were generated from the same multivariate Gaussian distribution (Eq 4) where  $\Sigma$  was a rank-one covariance matrix and the mean response vector,  $\mu$ , corresponded to one of two stimulus conditions,  $a$  or  $b$ . Top: Scree plot of noise covariance. Bottom: Cross-validated discriminability,  $d^2$ , between  $a$  and  $b$  computed with full-rank data and with *dDR* using one (*dDR*<sub>1</sub>), two (*dDR*<sub>2</sub>) or three (*dDR*<sub>3</sub>) noise dimensions, as a function of sample size. Mean  $d^2$  across all 100 surrogate datasets is shown here. For  $k \gg N$ , the *dDR* results converge to the asymptotic value of the full-rank  $d^2$ . However, even for small  $k$ , the *dDR* analyses estimates are much more accurate than the full-rank approach. **b.** Same as in (a), but for two-dimensional noise covariance data. In this case, *dDR*<sub>2</sub> captures the second noise dimension and outperforms the standard 1-D approach (*dDR*<sub>1</sub>). **c.** Same as in (a) and (b), but for  $1/n$  noise covariance.

<https://doi.org/10.1371/journal.pone.0271136.g005>

until  $k$  becomes large and they diminish with increasing noise dimensions—the improvement of *dDR*<sub>2</sub> over *dDR*<sub>1</sub> is larger than that of *dDR*<sub>3</sub> to *dDR*<sub>2</sub> (Fig 5b and 5c). This is because subsequent noise dimensions are, by definition, lower variance and therefore more difficult to estimate reliably from limited sample sizes.

### *dDR* recovers more decoding information than standard principal component analysis

One popular method for dimensionality reduction of neural data is principal component analysis (PCA) [33]. Generally speaking, PCA can be implemented on neural data in one of two ways: application to single trial spike counts PCA or trial-averaged spike counts PCA. In the single trial approach (*stPCA*), principal components are measured across all single trials and all experimental conditions. The resulting PCs capture variance both across trials and across *e.g.* stimulus conditions. In trial-averaged PCA (*taPCA*), single trial responses are first averaged per experimental condition and PCs are measured over the resulting  $N$ -neuron  $\times$   $S$ -condition spike count matrix. In this case, for different stimulus conditions, the PCs specifically capture variance of stimulus-evoked activity rather than trial-trial variability, making it a more logical choice for many decoding applications. In the case of  $S = 2$ , as we have outlined above for the *dDR* illustration (Fig 4), *taPCA* is equivalent to  $\Delta\mu$ , the first *dDR* dimension. Thus, *dDR* can

roughly be thought of as a way to combine *taPCA* and *stPCA*—*taPCA* identifies the *signal* dimension and *stPCA* identifies the *noise* dimension(s).

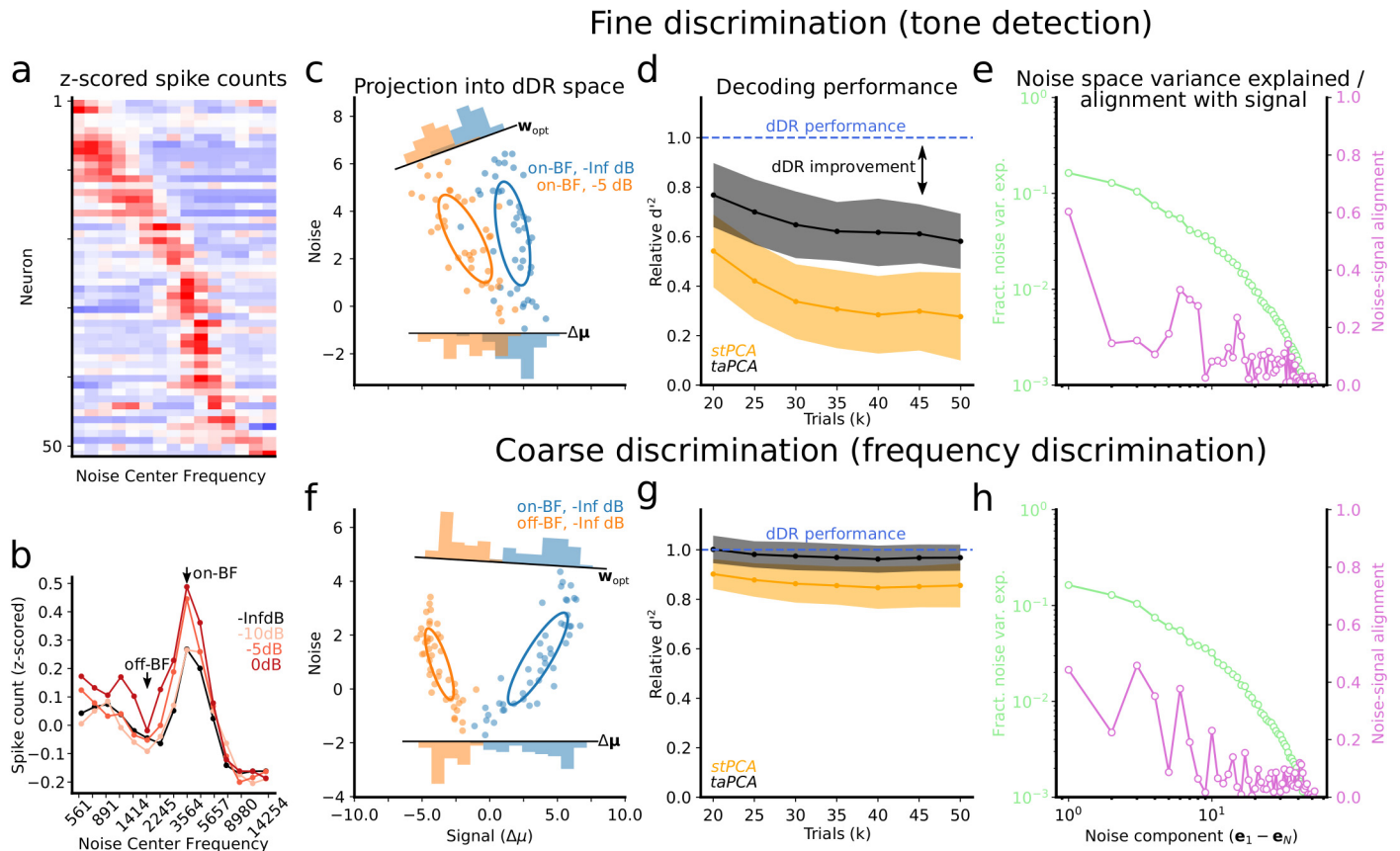
To demonstrate the relative decoding performance achieved using each method, we applied each to a dataset collected from primary auditory cortex in an awake, passively listening ferret.  $N = 52$  neurons were recorded simultaneously using a 64-channel laminar probe [22] as in [19, 20, 37]. Auditory stimuli consisting of narrowband (0.3 octave bandwidth) noise bursts were presented alone (-Inf dB) or with a pure tone embedded at varying SNRs (0 dB, -5 dB, -10 dB) in the hemifield contralateral to the recording site (see Experimental Methods). Each stimulus was repeated 50 times. The neural response to each stimulus was defined as the total number of spikes detected during the 300 ms stimulus duration for each neuron. For *stPCA* and *dDR*, we selected only the top  $m = 2$  total dimensions, and for *taPCA*, we selected the single dimension,  $\Delta\mu$ , that exists for  $S = 2$ . This dataset allowed us to investigate how each dimensionality reduction method performs for two distinct, behaviorally relevant neural decoding questions: One, how well can neural activity perform fine discriminations (*tone-in-noise detection*), discriminating noise alone vs. noise with tone? Two, how well can it perform coarse discriminations (*frequency discrimination*), discriminating noise centered at frequency A vs. noise at frequency B?

The A1 dataset displayed a range of frequency tuning (Fig 6a), with the majority of units tuned to  $\approx 3.5$  kHz. We therefore defined this as the best frequency of the recording site (on-BF, Fig 6b). For *tone detection*, we measured discriminability ( $d'^2$ , Eq 1) between on-BF noise alone (on-BF, -Inf dB) and on-BF noise plus tone (on-BF, -5 dB), which each drove similar sensory responses (Fig 6b and 6c). For *frequency discrimination*, we measured discriminability between the neural responses to on-BF noise and off-BF noise, where off-BF was defined as  $\approx 1$  octave away from BF, and drove a very different population response (Fig 6b and 6f). In both cases, *taPCA* and *dDR* outperformed *stPCA* (Fig 6d and 6g). This first result is unsurprising due to the fact that *stPCA* is the only method not explicitly designed to capture variability in the sensory response. The top PCs are dominated by dimensions of trial-trial variability that do not necessarily contain stimulus information and thus underestimate  $d'^2$  relative to the other two methods.

We also find that *dDR* consistently performs as well or better than *taPCA*. For the *tone detection* data, the sensory signal ( $\Delta\mu$ ) is small (*i.e.*, trial-averaged responses to the two stimuli were similar) and covariability is partly aligned with  $\Delta\mu$ . Under these conditions, *dDR* makes use of correlated activity to optimize the decoding axis ( $w_{opt}$ ) and improve discriminability. *taPCA*, on the other hand, has no information about these correlations and is therefore equivalent to projecting the single trial responses onto the *signal* axis,  $\Delta\mu$ . Thus, it underestimates  $d'^2$  (Fig 6c and 6d). In the *frequency discrimination* example,  $\Delta\mu$  is large. The covariability has similar magnitude to the previous example, but it is not aligned to the discrimination axis, and thus has no impact on  $w_{opt}$ . In this case, *dDR* and *taPCA* perform similarly (Fig 6f and 6g). These examples highlight that under behaviorally relevant conditions, *dDR* can offer a significant improvement over standard PCA, even with as few as 20 trials.

## Discussion

We have described a simple new method for robust decoding analysis of neural population data, decoding-based dimensionality reduction (*dDR*). This approach combines strategies for both trial-averaged PCA and single-trial PCA to identify important dimensions of population activity that govern neural coding accuracy. Using both simulated and real neural data, we demonstrated that the method performs robustly for neural decoding analysis in low experimental trial count regimes where the performance of full-rank methods break down. Across a



**Fig 6. *dDR* outperforms *PCA* for fine sensory discrimination.** **a.** Heatmap shows mean z-scored spike counts of  $N = 52$  simultaneously recorded units for 15 different narrowband noise bursts (0.3 octave bandwidth tiling 5 octaves,  $x$ -axis). Each row shows tuning for one unit, with red indicating higher firing rate response.  $x$ -axis (Noise Center Frequency of the sound stimulus) is shared with panel **b**. **b.** Population tuning curve for noise alone (black, data from panel **a**) and noise plus  $-10$ ,  $-5$ , and  $0$  dB tones (light to dark red), computed by averaging tuning curves across neurons. **c-e.** Decoding analysis for tone-in-noise detection. **c.** Scatter plot compares single trial responses to noise alone at best frequency (on-BF, blue) vs. noise +  $-5$  dB tone (orange), projected into *dDR* space. Ellipses show standard deviation across trials, marginal histograms show projection of data onto optimal decoding axis ( $w_{opt}$ ) or onto  $\Delta\mu$  (equivalent to performing trial-averaged *PCA*). **d.** Estimate of relative  $d^2$  as a function of sample size (number of trials,  $k$ ) using each dimensionality reduction method. For each data point, relative  $d^2$  was measured between *taPCA* vs. *dDR* and *stPCA* vs. *dDR*. This metric was averaged over 200 bootstrap samples of  $k$  trials. Shading indicates standard error. See [Methods](#) for details. **e.** Fraction variance explained by each noise component (green) computed by performing *PCA* on mean-centered single trial data. The alignment of each noise component with the signal axis is shown in purple. **f-h** Same as panels (c)-(e), for noise alone on-BF vs. noise alone off-BF (see panel **b**).

<https://doi.org/10.1371/journal.pone.0271136.g006>

range of behaviorally relevant stimulus conditions, *dDR* consistently performs as well or better than standard principal component analysis.

### Applications

*dDR* is designed to optimize the performance of linear decoding methods in situations where sample sizes are small. This is often the case for neurophysiology data collected from behaving animals, where the number of stimulus and/or behavior conditions are fundamentally limited by task performance. In these situations, using full-rank decoding methods is unfeasible as it leads to dramatic overfitting and unreliable performance [12]. Dimensionality reduction methods, such as *PCA*, can be used to mitigate overfitting issues. However, the correct implementation of *PCA* in neural data is often ambiguous, and multiple different approaches to dimensionality reduction have been proposed [33]. We suggest *dDR* as a simple, standardized alternative that captures the strengths of different *PCA* approaches. Unlike conventional *PCA*,

the *signal* and *noise* axes that comprise the *dDR* space have clear interpretations with respect to neural decoding. Importantly, *dDR* components explicitly preserve stimulus-independent population covariability. In addition to being important for overall information coding, this covariability is known to depend on behavior state [15, 16, 30, 34, 38] and stimulus condition [31, 39–41]. Therefore, approaches that do not preserve these dynamics, such as trial-averaged *PCA*, may not accurately characterize how information coding changes across varying behavior and/or stimulus conditions.

### Interpretability and visualization

A key benefit of *dDR* is that the axes making up the *dDR* subspace are easily interpretable: The first axis (*signal*) represents the dimension with maximal information about the difference in evoked activity between the two conditions to be decoded, and the second (*noise*) axis captures the largest mode of condition-independent population covariability in the data. Therefore, within the *dDR* framework it is straightforward to investigate how this covariability interacts with discrimination, an important question for neural information coding. Further, standard *dDR* (with a single noise dimension) can be used to easily visualize high-dimensional population data, as in Fig 6. For methods like *PCA*, it can be difficult to dissociate signal and noise dimensions, as the individual principal components can represent an ambiguous mix of task conditions, stimulus conditions, and trial-trial variability [42]. Moreover, with *PCA* the number of total dimensions is typically selected based on their cumulative variance explained, rather than by selecting the dimensions that are of interest for decoding, as in *dDR*.

### Extensions

**Latent variable estimation.** *dDR* makes the assumption that latent sources of low-dimensional neural variability can be captured using simple, linear methods, such as *PCA*. While these methods often seem to recover meaningful dimensions of neural variability [16], a growing body of work is investigating new, alternative methods for estimating these latent dynamics [15, 17, 43–45], and this work will continue to lead to important insights about the nature of shared variability in neural populations.

We suggest that *dDR* can be extended to incorporate these new methods. For example, rather than defining *dDR* on a strictly per decoding pair basis, a global noise axis could be identified across all experimental conditions using a custom latent variable method. This could then be applied to the decoding-based dimensionality reduction such that the resulting *dDR* space explicitly preserves activity in this latent space to investigate how it interacts with coding.

**Incorporating additional *dDR* dimensions.** In this work we have described *dDR* primarily as a transformation from  $N$ -dimensions to two dimensions, *signal* and *noise*, with the exception of Fig 5. In our code repository, <https://github.com/crheller/dDR>, we include examples that demonstrate how the *dDR* method can be extended to include additional dimensions. However, as discussed in the main text, it is important to remember that estimates of neural variability beyond the first principal component may become unreliable as variance along these dimensions gets progressively smaller, especially in low trial regimes. In short, while information may be contained in dimensions  $> m = 2$ , caution should be used to ensure that these dimensions can be estimated reliably.

### Related methods for dimensionality reduction and neural decoding

A growing number of techniques exist for performing dimensionality reduction on neural data [33]. It is outside the scope of this work to provide an exhaustive review of these methods, however, it is useful to highlight a couple of methods that share similarities with *dDR*.

Standard principal component analysis (*PCA*) remains the most commonly applied method for dimensionality reduction of neural data. Indeed, many other methods, (*e.g.* k-means clustering, non-negative matrix factorization etc.), can simply be viewed specially constrained versions of *PCA* [46]. In the case of *dDR*, the key distinction from *PCA* is that *dDR* explicitly preserves information about both trial-trial neural variability and mean, *e.g.* stimulus evoked, activity. *PCA*, on the other hand, is an entirely unsupervised method. Therefore, individual principal components often contain a mixture of trial-trial variability and mean activity, making their interpretation challenging. Furthermore, performing *PCA* can lead to sub-optimal neural decoding, as we demonstrate above.

Recently, Kobak et al. developed a powerful method called demixed *PCA* (*dPCA*) which produces interpretable low-dimensional representations of neural population data [42]. While this work shares some conceptual similarities with *dDR*, namely that both measure interpretable, low-dimensional representations of neural data, their applications are distinct. With *dPCA*, the idea is to produce components that allow accurate decoding of *e.g.* stimulus condition and also maintain a faithful representation of the underlying neural state space geometry. In the context of optimal decoding, these two aims can sometimes be at odds—the best geometrical representation of the data does not necessarily lead to optimal decodability. This trade off is illustrated nicely in their manuscript [42].

Unlike *dPCA*, *dDR*, only seeks to maximize information that can be used for decoding. Therefore, *dDR* can be thought of as a preprocessing step that should be applied prior to performing standard decoding methods, such as linear discriminant analysis (*LDA*). Further, *dDR* is designed for optimal decoding of only two experimental conditions at a time. *dPCA* is not restricted in this pairwise way. Therefore, *dPCA* is useful when an interpretable, constant low-*D* space across many different experimental conditions is desired, while *dDR* should be used when optimal decoding is the goal.

## Code availability

We provide Python code for *dDR* which can be downloaded and installed by following the instructions at <https://github.com/crheller/dDR>. We also include a short demo notebook that highlights the basic work flow and implementation of the method to simulated data. All code used to generate the figures in this manuscript is available in the repository.

## Supporting information

**S1 Appendix.**  
(PDF)

## Author Contributions

**Conceptualization:** Charles R. Heller, Stephen V. David.

**Data curation:** Charles R. Heller.

**Formal analysis:** Charles R. Heller.

**Funding acquisition:** Charles R. Heller, Stephen V. David.

**Investigation:** Charles R. Heller.

**Methodology:** Charles R. Heller.

**Project administration:** Stephen V. David.

**Resources:** Stephen V. David.

**Software:** Charles R. Heller.

**Supervision:** Stephen V. David.

**Validation:** Stephen V. David.

**Visualization:** Charles R. Heller.

**Writing – original draft:** Charles R. Heller.

**Writing – review & editing:** Charles R. Heller, Stephen V. David.

## References

1. Britten KH, Shadlen MN, Newsome WT, Movshon JA. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*. 1992; 12(12):4745–4765. <https://doi.org/10.1523/JNEUROSCI.12-12-04745.1992> PMID: 1464765
2. Zohary E, Shadlen MN, Newsome WT. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature*. 1994; 370(6485):140–143. <https://doi.org/10.1038/370140a0> PMID: 8022482
3. Shadlen MN, Newsome WT. The Variable Discharge of Cortical Neurons: Implications for Connectivity, Computation, and Information Coding. *Journal of Neuroscience*. 1998; 18(10):3870–3896. <https://doi.org/10.1523/JNEUROSCI.18-10-03870.1998> PMID: 9570816
4. Abbott LF, Dayan P. The Effect of Correlated Variability on the Accuracy of a Population Code. *Neural Computation*. 1999; 11(1):91–101. <https://doi.org/10.1162/089976699300016827> PMID: 9950724
5. Dayan P, Abbott LF. *Theoretical neuroscience: computational and mathematical modeling of neural systems*. Computational neuroscience. Cambridge, Mass: Massachusetts Institute of Technology Press; 2001.
6. Averbeck BB, Lee D. Effects of noise correlations on information encoding and decoding. *Journal of Neurophysiology*. 2006; 95(6):3633–3644. <https://doi.org/10.1152/jn.00919.2005> PMID: 16554512
7. Averbeck BB, Latham PE, Pouget A. Neural correlations, population coding and computation. *Nature Reviews Neuroscience*. 2006; 7(5):358–366. <https://doi.org/10.1038/nrn1888> PMID: 16760916
8. Pitkow X, Liu S, Angelaki D, DeAngelis G, Pouget A. How Can Single Sensory Neurons Predict Behavior? *Neuron*. 2015; 87(2):411–423. <https://doi.org/10.1016/j.neuron.2015.06.033> PMID: 26182422
9. Bartolo R, Saunders RC, Mitz AR, Averbeck BB. Information limiting correlations in large neural populations. *The Journal of Neuroscience*. 2020; p. 2072–19. <https://doi.org/10.1523/JNEUROSCI.2072-19.2019> PMID: 31941667
10. Kafashan M, Jaffe A, Chettih SN, Nogueira R, Arandia-Romero I, Harvey CD, et al. Scaling of information in large neural populations reveals signatures of information-limiting correlations. *bioRxiv*. 2020; p. 2020.01.10.902171.
11. Romyantsev OI, Lecoq JA, Hernandez O, Zhang Y, Savall J, Chrapkiewicz R, et al. Fundamental bounds on the fidelity of sensory cortical coding. *Nature*. 2020; <https://doi.org/10.1038/s41586-020-2130-2> PMID: 32238928
12. Kanitscheider I, Coen-Cagli R, Kohn A, Pouget A. Measuring Fisher Information Accurately in Correlated Neural Populations. *PLOS Computational Biology*. 2015; 11(6):e1004218. <https://doi.org/10.1371/journal.pcbi.1004218> PMID: 26030735
13. Stevenson IH, Kording KP. How advances in neural recording affect data analysis. *Nature Neuroscience*. 2011; 14(2):139–142. <https://doi.org/10.1038/nn.2731> PMID: 21270781
14. Cowley BR, Snyder AC, Acar K, Williamson RC, Yu BM, Smith MA. Slow Drift of Neural Activity as a Signature of Impulsivity in Macaque Visual and Prefrontal Cortex. *Neuron*. 2020; 108(3):551–567.e8. <https://doi.org/10.1016/j.neuron.2020.07.021> PMID: 32810433
15. Rabinowitz NC, Goris RL, Cohen M, Simoncelli EP. Attention stabilizes the shared gain of V4 populations. *eLife*. 2015; 4:e08998. <https://doi.org/10.7554/eLife.08998> PMID: 26523390
16. Ni AM, Ruff DA, Alberts JJ, Symmonds J, Cohen MR. Learning and attention reveal a general relationship between population activity and behavior. *Science*. 2018; 359(6374):463–465. <https://doi.org/10.1126/science.aao0284> PMID: 29371470

17. Ecker A, Berens P, Cotton RJ, Subramaniyan M, Denfield G, Cadwell C, et al. State Dependence of Noise Correlations in Macaque Primary Visual Cortex. *Neuron*. 2014; 82(1):235–248. <https://doi.org/10.1016/j.neuron.2014.02.006> PMID: 24698278
18. Slee SJ, David SV. Rapid Task-Related Plasticity of Spectrotemporal Receptive Fields in the Auditory Midbrain. *Journal of Neuroscience*. 2015; 35(38):13090–13102. <https://doi.org/10.1523/JNEUROSCI.1671-15.2015> PMID: 26400939
19. Heller CR, Schwartz ZP, Sadari D, David SV. Selective effects of arousal on population coding of natural sounds in auditory cortex. *bioRxiv*. 2020; p. 2020.08.31.276584.
20. Sadari D, Schwartz ZP, Heller CR, Pennington JR, David SV. Dissociation of task engagement and arousal effects in auditory cortex and midbrain. *eLife*. 2021; 10:e60153. <https://doi.org/10.7554/eLife.60153> PMID: 33570493
21. Bizley JK, Nodal FR, Nelken I, King AJ. Functional organization of ferret auditory cortex. *Cerebral Cortex (New York, NY: 1991)*. 2005; 15(10):1637–1653. PMID: 15703254
22. Du J, Blanche TJ, Harrison RR, Lester HA, Masmanidis SC. Multiplexed, High Density Electrophysiology with Nanofabricated Neural Probes. *PLOS ONE*. 2011; 6(10):e26204. <https://doi.org/10.1371/journal.pone.0026204> PMID: 22022568
23. Siegle JH, López AC, Patel YA, Abramov K, Ohayon S, Voigts J. Open Ephys: an open-source, plugin-based platform for multichannel electrophysiology. *Journal of Neural Engineering*. 2017; 14(4):045003. <https://doi.org/10.1088/1741-2552/aa5eea> PMID: 28169219
24. Efron B, Tibshirani R. Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statistical Science*. 1986; 1(1):54–75. <https://doi.org/10.1214/ss/1177013817>
25. Bickel PJ, Freedman DA. Some Asymptotic Theory for the Bootstrap. *The Annals of Statistics*. 1981; 9(6):1196–1217. <https://doi.org/10.1214/aos/1176345637>
26. Stringer C, Pachitariu M, Steinmetz N, Reddy CB, Carandini M, Harris KD. Spontaneous behaviors drive multidimensional, brainwide activity. *Science*. 2019; 364(6437):eaav7893. <https://doi.org/10.1126/science.aav7893> PMID: 31000656
27. Goris RLT, Movshon JA, Simoncelli EP. Partitioning neuronal variability. *Nature Neuroscience*. 2014; 17(6):858–865. <https://doi.org/10.1038/nn.3711> PMID: 24777419
28. Moreno-Bote R, Beck J, Kanitscheider I, Pitkow X, Latham P, Pouget A. Information-limiting correlations. *Nature Neuroscience*. 2014; 17(10):1410–1417. <https://doi.org/10.1038/nn.3807> PMID: 25195105
29. Green DM, Swets JA. Signal detection theory and psychophysics. Signal detection theory and psychophysics. Oxford, England: John Wiley; 1966.
30. Cohen MR, Maunsell JHR. Attention improves performance primarily by reducing interneuronal correlations. *Nature Neuroscience*. 2009; 12(12):1594–1600. <https://doi.org/10.1038/nn.2439> PMID: 19915566
31. Cohen MR, Kohn A. Measuring and interpreting neuronal correlations. *Nature Neuroscience*. 2011; 14(7):811–819. <https://doi.org/10.1038/nn.2842> PMID: 21709677
32. Ecker AS, Berens P, Keliris GA, Bethge M, Logothetis NK, Tolias AS. Decorrelated neuronal firing in cortical microcircuits. *Science (New York, NY)*. 2010; 327(5965):584–587. <https://doi.org/10.1126/science.1179867> PMID: 20110506
33. Cunningham JP, Yu BM. Dimensionality reduction for large-scale neural recordings. *Nature Neuroscience*. 2014; 17(11):1500–1509. <https://doi.org/10.1038/nn.3776> PMID: 25151264
34. Downer JD, Niwa M, Sutter ML. Task Engagement Selectively Modulates Neural Correlations in Primary Auditory Cortex. *Journal of Neuroscience*. 2015; 35(19):7565–7574. <https://doi.org/10.1523/JNEUROSCI.4094-14.2015> PMID: 25972181
35. Ni AM, Bowes BS, Ruff DA, Cohen MR. Methylphenidate as a causal test of translational and basic neural coding hypotheses. *Proceedings of the National Academy of Sciences*. 2022; 119(17):e2120529119. <https://doi.org/10.1073/pnas.2120529119> PMID: 35467980
36. Stringer C, Pachitariu M, Steinmetz N, Carandini M, Harris KD. High-dimensional geometry of population responses in visual cortex. *Nature*. 2019; 571(7765):361–365. <https://doi.org/10.1038/s41586-019-1346-5> PMID: 31243367
37. Pennington J, David S. Complementary effects of adaptation and gain control on sound encoding in primary auditory cortex. *Neuroscience*; 2020. Available from: <http://biorxiv.org/lookup/doi/10.1101/2020.01.14.905000>. <https://doi.org/10.1523/ENEURO.0205-20.2020> PMID: 33109632
38. Valente M, Pica G, Runyan CA, Morcos AS, Harvey CD, Panzeri S. Correlations enhance the behavioral readout of neural population activity in association cortex. *Neuroscience*; 2020. Available from:



- <http://biorxiv.org/lookup/doi/10.1101/2020.04.03.024133>. <https://doi.org/10.1038/s41593-021-00845-1> PMID: 33986549
39. Zylberberg J, Cafaro J, Turner M, Shea-Brown E, Rieke F. Direction-Selective Circuits Shape Noise to Ensure a Precise Population Code. *Neuron*. 2016; 89(2):369–383. <https://doi.org/10.1016/j.neuron.2015.11.019> PMID: 26796691
  40. Franke F, Fiscella M, Sevelev M, Roska B, Hierlemann A, da Silveira R Azeredo. Structures of Neural Correlation and How They Favor Coding. *Neuron*. 2016; 89(2):409–422. <https://doi.org/10.1016/j.neuron.2015.12.037> PMID: 26796692
  41. Ruff DA, Cohen MR. Stimulus Dependence of Correlated Variability across Cortical Areas. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*. 2016; 36(28):7546–7556. <https://doi.org/10.1523/JNEUROSCI.0504-16.2016> PMID: 27413163
  42. Kobak D, Brendel W, Constantinidis C, Feierstein CE, Kepecs A, Mainen ZF, et al. Demixed principal component analysis of neural population data. *eLife*. 2016; 5:e10989. <https://doi.org/10.7554/eLife.10989> PMID: 27067378
  43. Whiteway MR, Averbeck B, Butts DA. A latent variable approach to decoding neural population activity. *bioRxiv*. 2020; p. 2020.01.06.896423.
  44. Yu BM, Cunningham JP, Santhanam G, Ryu SI, Shenoy KV, Sahani M. Gaussian-Process Factor Analysis for Low-Dimensional Single-Trial Analysis of Neural Population Activity. *Journal of Neurophysiology*. 2009; 102(1):614–635. <https://doi.org/10.1152/jn.90941.2008> PMID: 19357332
  45. Aoi MC, Mante V, Pillow JW. Prefrontal cortex exhibits multidimensional dynamic encoding during decision-making. *Nature Neuroscience*. 2020; 23(11):1410–1420. <https://doi.org/10.1038/s41593-020-0696-5> PMID: 33020653
  46. Udell M, Horn C, Zadeh R, Boyd S. Generalized Low Rank Models. *Foundations and Trends in Machine Learning*. 2016; 9(1):1–118. <https://doi.org/10.1561/22000000055> PMID: 26776181