Homeostatic synaptic scaling optimizes learning in network models of neural population codes

Jonathan Mayzel and Elad Schneidman

Department of Brain Sciences, Weizmann Institute of Science, Rehovot 76100, Israel

Studying and understanding the code of large neural populations hinge on learning accurate statistical models of population activity. A new class of such models, based on learning to weigh sparse nonlinear Random Projections (RP) of the population, was recently shown to be highly accurate, efficient, and scalable. Moreover, RP models have a clear biologically-plausible implementation as a shallow neural circuit. Here we extend these models and present RP models that are learned by optimizing the randomly selected sparse projections. This "reshaping" of projections is akin to changing synaptic connections in the corresponding neural circuit model. When we applied these Reshaped RP models to recordings of tens of cortical neurons from behaving monkeys, we found them to be more accurate and efficient than the previous class of RP models and on par with backpropagation models. Our exploration of the effect of adding biological features to these circuit models revealed that learning reshaped RP models with homeostatic synaptic scaling yields even more efficient and accurate models. We further show that homeostatic reshaped RP models, which rely on sparse and random connectivity, are superior to fully connected network models. Our results thus suggest a key functional role for homeostatic scaling in neural circuits, beyond regulating network activity, namely – optimizing performance and efficiency.

Population Codes | **Neural Networks** | **Homeostatic Plasticity** | **Random Projections** | **Efficient Coding** | **Learning**

Correspondence: ***elad.schneidman@weizmann.ac.il***

Introduction

The potential "vocabulary" of the spiking patterns of a population of neurons scales exponentially with the size of the population we may consider. Therefore, mapping the rules of neural population codes and their semantic organization cannot rely on direct sampling of the vocabulary for more than a handful of neurons. Moreover, the stochastic nature of neural activity implies that the characterization of neural codes must rely on probability distributions over population activity patterns. Thus, to describe and analyze the structure and content of the code with which neural circuits respond to stimuli or direct action – we need to learn statistical models of their activity. Such models have been used to study the design of neural population codes and their content in different systems: Models of the directional coupling between neurons, such as Generalized Linear Models, have been used to replicate the stimulus-dependent rates of populations of tens of neurons (1–4). Maximum entropy models have accurately captured the joint activity patterns of more than 100 neurons, using simple statistical features of the population, like firing rates, pairwise correlations, synchrony, and other low-order statistics (5–13). They have further been used to explore the semantic organization of the code (14, 15). Auto-encoders have also been employed to replicate the detailed structure of population activity – yielding generative models that can be used to study the code, but their design is difficult to interpret (16–18). Importantly, scaling these models to hundreds of neurons is computationally challenging (9, 19, 20) and is, therefore, a key hurdle in modeling large neural systems.

While statistical models are invaluable for describing and studying neural codes, it is not immediately clear whether the brain relies on such models or implements them when representing or processing information (21, 22). In particular, much of the analysis of neural codes has focused on decoding population activity, typically using simple decoders (2, 23–27), or metrics over the structure of population activity patterns (14, 28, 29). Yet, if neural circuits implement such statistical models, and in particular, ones that compute the likelihood of their inputs – it would be instrumental for carrying Bayesian computation and decision making (30–32). Moreover, learning such network models would be of key interest, both as a potential way to improve learning in artificial neural networks using biological features, and as a way for biological neural networks to implement efficient learning and overcome the credit assignment problem (33–38).

Both structured architectural features of neural circuits and random connectivity patterns have been suggested to enable or shape the computation carried out by neural circuits (30, 39–43). In particular, these computations rely on the nature of synaptic connectivity and the coupling between synapses in terms of how they change during learning. Competition mechanisms between synapses or other regularization mechanisms have also been suggested to be important components of computation and learning, for example, in cortical circuits (44, 45). One such mechanism is the homeostatic scaling of synaptic plasticity, which has been observed *in vitro* and *in vivo* at the level of incoming synapses to a neuron and outgoing ones (46–49). This mechanism has been commonly attributed to the regulation of firing rates, while its computational implications remain mostly unclear, but of interest computationally and mechanistically (50–54). A related computational feature has been presented by network models that include divisive normalization, which has been suggested to be a key component of computations performed by cortical circuits (55).

Here, we explore biologically-inspired extensions of a new family of statistical models for large neural population codes. Adding biological features and constraints to population models enabled us to ask how to improve the models, and how the design of real neural circuits could allow them to implement such models. Specifically, we expand the Random Projections (RP) model (30), which was shown to be

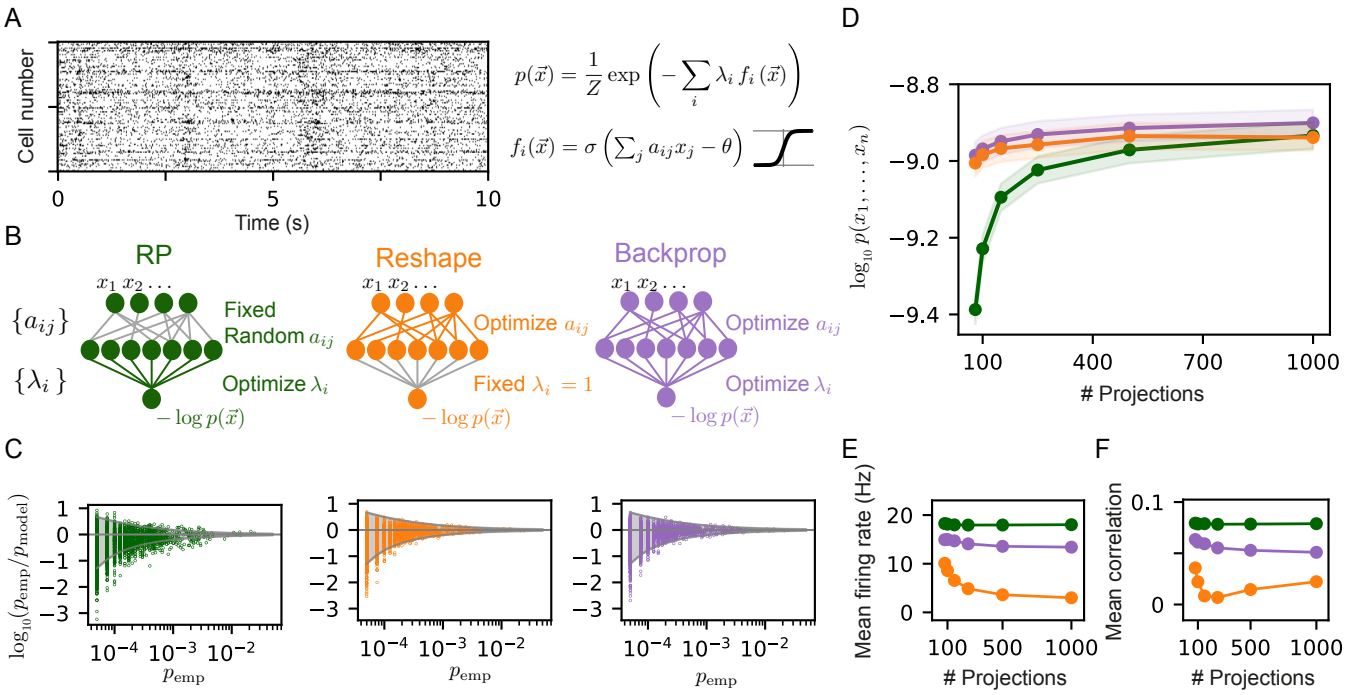


Fig. 1. Reshaped Random Projections models outperform Random Projections models. (A) A short segment of the spiking activity of 100 cortical neurons used for the analysis and comparison of different statistical models of population activity (see Methods). (B) Schematics of the neural circuits that implement the different random projections models we compared: The "standard" Random Projections model, where the coefficients a_{ij} that define the projections are randomly selected and fixed whereas the factors λ_i are learned (see text). The Reshaped Random Projections model, where the coefficients a_{ij} that define the projections are tuned and the factors λ_is are fixed. The backpropagation model, where we tune both the a_{ij}s and λ_is. (C) The predicted probability of individual populations' activity patterns as a function of the observed probability for RP, Reshaped and Backpropagation models. Gray funnels denote 99% confidence interval. (D) Average performance of models of the three classes is shown as a function of the number of projections, measured by log-likelihood, over 100 sets of randomly selected groups of 50 neurons. Reshaped models outperform RP models and are on par with backpropagation; the shaded area denotes the standard error over 100 models. (E-F) Mean firing rates of projection neurons and mean correlation between projections. Reshaped models show lower correlations and lower firing rates compared to RP and backpropagation models. Standard errors are smaller than marker's size, hence invisible.

highly accurate at the level of more than 100 neurons in different systems, and which relies on a random set of sparse and non-linear features of the population. Importantly, in addition to being accurate and requiring little amounts of training data, RP models can be readily implemented by a simple neural circuit model – suggesting how real neural circuits can learn a statistical model of their own inputs. We then asked whether we can make these models better by "reshaping" the randomly chosen projections, and show that we can indeed achieve highly accurate models using significantly fewer projections. Using biologically-plausible circuits that would implement RP models, we further asked how would adding biological characteristics to the reshaping of projections affect the model. We found that reshaping that incorporates homeostatic scaling results in more accurate models that are also more efficient. Thus, we present a new class of accurate and efficient statistical models for large neural population codes that demonstrates a clear computational benefit of homeostatic synaptic scaling and its potential role in biological neural networks and artificial ones.

Results

The Random Projections (RP) model is a class of highly accurate, scalabale, and efficient statistical models of the joint activity patterns of large populations of neurons (30, 31). These models are based on random and sparse nonlinear functions, or "projections", of the population: Given a recording of the spiking activity of a neuronal population, the model gives a probability distribution over a discrete time representation of the activity (e.g., 10-20 ms bins), using a set of random non-linear functions of the population activity,

$$f_i(\vec{x}) = \sigma\left(\sum a_{ij}x_j - \theta_j\right), \qquad \textbf{(1)}$$

where a_{ij} are randomly sampled coefficients such that most of them for any i are zero (i.e., the set is sparse), θ_j are thresholds, and $\sigma(\cdot)$ are nonlinear functions, (e.g., the Heaviside step function). The RP model is defined as the maximum entropy distribution $p(\vec{x})$ (56), which is consistent with the observed average values of the random projections $\langle f_i\rangle_p = \langle f_i\rangle_{data}$ (See Methods). This distribution is unique, given by

$$p_{RP}(\vec{x}) = \frac{1}{Z}\exp\left(-\sum_i \lambda_i f_i(\vec{x})\right), \qquad \textbf{(2)}$$

where λ_i are Lagrange multipliers, and Z is a normalization factor or the "partition function", which can be found numerically. When applied to cortical data from multiple areas (see, e.g., Figure 1A), this model proved to be highly accurate using small amounts of training data (30). Importantly, unlike

many other statistical models of population activity, RP models have a simple, biologically plausible neural circuit that can implement them. Figure 1B shows such a feed-forward circuit with one intermediate layer and an output neuron, where the random coefficients of the sparse projections, a_{ij}, are the synaptic weights connecting the input neurons $\vec{x}$ to an intermediate layer of neurons f_i. Each intermediate neuron implements one projection of the input population. The Lagrange multipliers, λ_i, are the synaptic weights connecting the intermediate layer to the output neuron, whose membrane potential or output gives the log-likelihood of the activity pattern of $\vec{x}$, up to a normalization factor.

The model in eq. 2 harbors a duality between the projections, f_i, and their coefficients, λ_i: In the maximum entropy formalism of the RP model, the projections are randomly sampled and then fixed, and their corresponding weights, λ_i's, are tuned to maximize the entropy and satisfy the constraints. Alternatively, we consider the case of training the model by keeping the λ_i's fixed and changing or tuning the projections f_i to maximize the likelihood. In the corresponding neural circuit, this would imply that we would learn a circuit that implements the statistical model by training the sparse set of synaptic connections, a_{ij}, which define the projections, instead of training the synaptic weights, λ_i.

Notably, a variant of the RP model in which projections that were weighted by a low value of λ_i are pruned and replaced with new random projections proved to be more accurate than the original RP model, while using fewer projections (30). This procedure of pruning and replacement is a crude form of learning of the model through changing the projections, and finding more efficient ones. We asked here whether instead of the heuristic pruning and replacement, we can directly learn more accurate and efficient models by tuning the projections.

Reshaping random projections give more accurate and compact models. We, therefore, learned a new class of statistical models for populations of tens of cortical neurons from the prefrontal cortex of monkeys performing a visual classification task (57). These statistical models use random projections and are learned by tuning the projections: Given an initial draw of sparse projections, the random weights that define the projections, a_{ij}, are changed to maximize the likelihood of the model:

$$\Delta a_{ij} = \eta \lambda_i \left(\left\langle \frac{\partial \sigma(\vec{x})}{\partial a_{ij}} \right\rangle_p - \left\langle \frac{\partial \sigma(\vec{x})}{\partial a_{ij}} \right\rangle_{data} \right), \quad \textbf{(3)}$$

where η is the learning rate. Unlike the RP model presented in (30), here we used a sigmoid function for the nonlinearity of the projections,

$$\sigma(x) = \frac{1}{1+e^{-\beta x}}, \quad \textbf{(4)}$$

where β sets the slope of the sigmoid. In this formulation, the model ranges from an independent model of the population for $\beta \to 0$, to the original RP model (30) for $\beta \to \infty$. The rule for changing the projections (eq. 3) means that the set of inputs to each projection neuron is retained, but their relative weights are changed, and so the projections are "reshaped". We focus henceforth on the case of all $\lambda_i = 1$, which we term the Reshaped Random Projections (Reshaped RP) model, which is given by

$$p_{Reshape}(\vec{x}) = \frac{1}{Z} \exp\left(-\sum_i f_i(\vec{x}) \right). \quad \textbf{(5)}$$

We compared between the RP and the Reshaped RP models by quantifying their performance on the same set of initial projections. To that end, we first learned the RP model as in (30), using a Heaviside non-linearity for the projections, and compared them to RP models that used a sigmoid non-linearity (where both models used the same set of random projections). The latter models were found to be more accurate (see Supplementary Figure 1). We then learned a Reshaped RP model in which we optimize the same initial projections while keeping $\lambda_i = 1$. We note that while in its maximum entropy formulation, the RP model is the unique solution to a convex optimization problem, the Reshaped RP models are not guaranteed to reach a global optimum. We also considered a third class of models, ones in which the projections and the Lagrange multipliers λ_i are optimized simultaneously, similar to backpropagation-based learning used to train feed-forward networks (see Methods). Figure 1C shows an example of the ability of the sigmoid RP models, Reshaped Random Projections models, and backpropagation-based models to predict the probability of individual activity patterns for one group of 20 neurons, recorded from the cortex of behaving monkeys (57). The activity patterns are predicted by the reshaped RP model to an accuracy that is within the sampling noise level (denoted by the 99% confidence interval funnel), and is similar to the performance of the full backpropagation model. The standard RP model, in comparison, has many more patterns that are outside the 99% confidence interval funnel. We quantified the performance of the three classes of models by calculating the mean log-likelihood of the models over 100 groups of 50 neurons on held out datasets, as a function of the number of projections that we used (Fig. 1D). The reshaped models outperform the RP ones for a low number of projections, whereas when many projections are used, the performances of all three models converge to a similar value.

To compare the "mechanistic" nature of these different models, we calculated the mean correlation between the projections within each model class, and the average values of each projection (where the average is over the population activity patterns), which correspond to the mean firing rates of the neurons in the intermediate layer. Interestingly, the firing rates of the neurons in the intermediate layer are considerably lower for the reshaped models, and this sparseness in activity becomes more pronounced as a function of the number of projections (Figure 1E). We further find that the correlations between the projections in the reshaped models are considerably lower compared to RP and backpropagation models (Figure 1F).

Thus, the reshaped projection models suggest a way to learn more accurate models of population activity, by the tuning

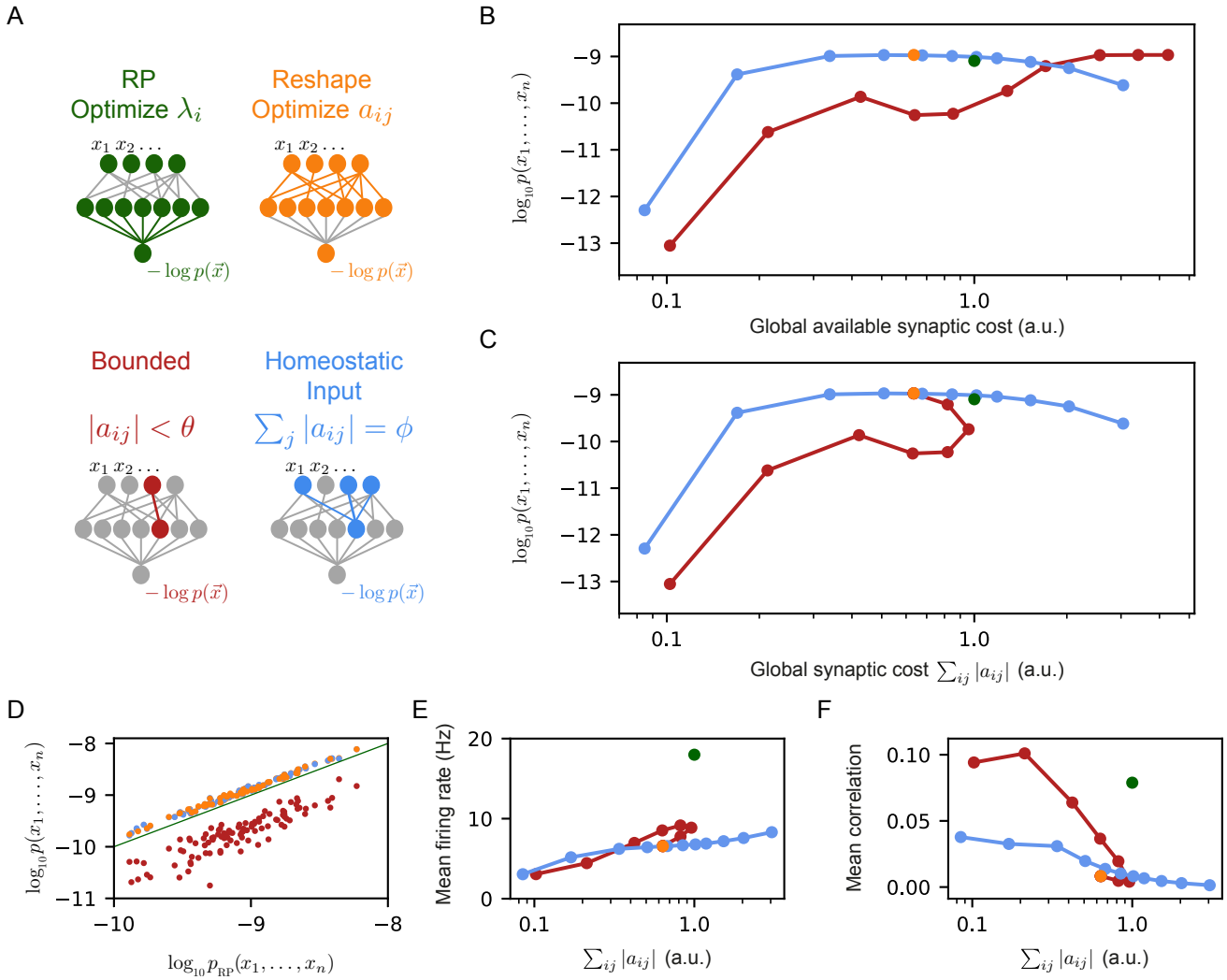


Fig. 2. Homeostatic Reshaped RP models outperform RP models and bounded RP models. (A) Schematic drawing of the different models we studied: standard RP model, unconstrained reshaped model, and two types of constrained reshaped models: Bounded models in which each synapse separately obeys $\sum_{ij} |a_{ij}| < \theta$ during learning, and Homeostatic input reshaped models, where we fix the total synaptic weight of incoming synapses $\sum_{ij} |a_{ij}| = \phi$. (B) The mean log-likelihood of the models is shown as a function of the total available budget. The Homeostatic input reshaped RP models give optimal results for a wide range of values of available synaptic budget, outperforming the bounded models and the RP model. (C) The mean log-likelihood of the models is shown as a function of the total used budget. Aside from the bounded models, all other models are the same as in (B) by construction. For high available budget values, bounded models show better performance while utilizing a lower synaptic budget, similar to the unconstrained reshape model. (D) Comparison of the performance of 100 individual examples of each model class and their corresponding RP model, where all models relied on the same set of initial projections. Homeostatic input models outperformed the RP models in all cases (all points are above the diagonal), while all bounded models were worse (points below the diagonal). (E-F) The mean correlation and firing rates of projections as a function of the model's cost. Homeostatic models show low correlations and mean firing rates, similar to unconstrained reshaped models. Note that in panels B, C, E, and F, the standard errors are smaller than the marker size, and are therefore invisible.

of projections. These models are also more efficient, requiring fewer projections, which have lower firing rates (i.e., reshaped projections use fewer spikes), and are less correlated. Given their accuracy and efficiency, we next asked how adding biological features or constraints to a Reshaped RP circuit may affect its performance and efficiency.

Homeostatic reshaping of random projections gives more accurate and synaptically efficient models. We next studied the effect of adding two classes of biological features or constraints to the model, on the performance and nature of the Reshaped RP circuit model. The first constraint is based on the biophysical limits on individual synapses, and so we bound the maximal strength of individual synapses such that the strength of all synaptic weights are smaller than a "ceiling" value: $|a_{ij}| < \theta$. The other is a homeostatic normalization of the synaptic weights, following the synaptic rescaling that has been observed experimentally (49). We consider multiple homeostatic mechanisms of this kind later on, but begin here with fixing the total sum of the incoming synaptic strength of each projection such that $\sum_j |a_{ij}| = \phi$. Thus, when the strength of one synapse increases (decreases), the strength of the rest of the incoming synapses decreases (increases) such that the total synaptic weight incoming into the projection is kept constant.

To compare the effect of these constraints, we used the same set of initial random projections, and then learn by reshaping

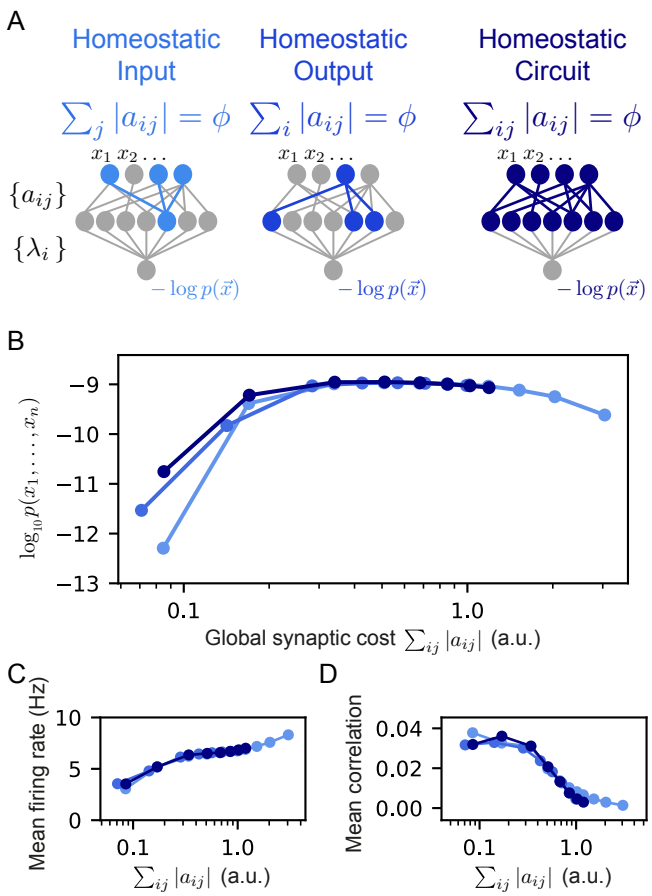


Fig. 3. Homeostatic model variants show similar results. (A) Schematic drawings of the different homeostatic models we compared: Homeostatic input models in which we fixed the total synaptic weight of the incoming synapses $\sum_j |a_{ij}| = \phi$; Homeostatic output models in which we fixed the total synaptic weight of the outgoing synapses $\sum_i |a_{ij}| = \phi$; and Homeostatic circuit models in which we fixed the total synaptic weight of the whole synaptic circuit $\sum_{ij} |a_{ij}| = \phi$. (B) The mean log-likelihood of models, shown as a function of the total used budget. All three homeostatic model variants show a similar behavior. (C) and (D) Mean correlation between the projection neurons and firing rates as a function of the model cost. Again, all three homeostatic constraints show a similar behavior. Note that in panels B, C, and D, the standard errors over 100 models are smaller than the marker's size and are, therefore, invisible.

them, each time with a different value of their corresponding parameters, ϕ or θ. We estimated the likelihood of each of the models on 100 groups of 50 neurons, over 100 random sets of 150 projections. To quantify the "synaptic budget" of each model, we measured the total sum of the absolute values of synaptic weights available to each model in units of the total synaptic strength of the initial set of projections (this is equivalent to defining the total sum of the synaptic weights of the initial set of projections as "1", and then measuring total synaptic weights in these units). For the models with bounded synapses, the total available synaptic budget is given by the number of synapses times θ, whereas for the homeostatic constraint, it equals ϕ times the number of projections in the model. Figure 2B shows the log-likelihood of each model class vs. the total available synaptic budget of the different models: For a wide range of synaptic budgets, the homeostatic models outperform the bounded models, and only for very high values of available synaptic budget, the performance of the bounded models is on par with the homeostatic models.

The differences between the homeostatic scaling models and the bounded synaptic strength models are further reflected in Figure 2C, where we show the performance of each model class as a function of the total sum of synaptic weights that is used by the model at the end of the training $\sum_{ij} |a_{ij}|$. We note that the curve of the homeostatic model is identical to the one from Figure 2B by definition, but the curve of the bounded models shows that at a certain value of θ the sum of the synaptic weights starts to decrease and converges to that of the unconstrained reshaped model. The poor performance of the bounded models compared to that of the homeostatic ones suggests that the coupled changes in the synaptic weights improve learning. Specifically, during reshaping, the homeostatic models "redistribute" their budget and move synaptic "mass" from less important synapses to more important ones. This redistribution of resources results in accurate models even for relatively low values of synaptic weights – making them more efficient in terms of the total synaptic weight needed.

The dominance of the homeostatic learning over the bounded synaptic weights is clear not just for the average over models, but also at the level of the individual model. Figure 2D shows the performance of the homeostatic and bounded models that are initialized with the same set of random projection. As is evident, all the bounded constraint models are inferior to the RP ones, whereas all the homeostatic constraint models are superior to the RP models (and clearly all the homeostatic models are superior to the corresponding bounded models).

We further find that the mean firing rates of the reshaped projections, as well as the correlations between them, are lower in the homeostatic models compared to the bounded models (Figure 2E-F), making them more energetically efficient (in terms of spiking activity), consistent with efficient coding by decorrelated neural populations (58, 59).

We explored two other homeostatic synaptic scaling rules for the reshaping of projections (Fig. 3A). In the first, we fixed and normalized the outgoing synapses from each neuron, such that that $\sum_i |a_{ij}| = \phi$. In the second, we kept the total synaptic weight of the whole circuit fixed $\sum_{ij} |a_{ij}| = \phi$. Figure 3B shows that the performance of the models that use these different homeostatic mechanisms is surprisingly similar in terms of the model's likelihood over the test data, as well as the firing rates of the projection neurons (Fig. 3C), and correlations between them (Fig. 3D).

Having established the computational benefits and efficiency of reshaped projection models that rely on homeostatic scaling, we turned to ask how the connectivity itself, rather than the synaptic weights, may affect the performance of the models.

Homeostatic reshaping of random projections is on par with optimal projections and better than a fully connected network. The benefits of reshaping the projections raises the question of the importance of the nature of the specific random projections, that are then reshaped. We,

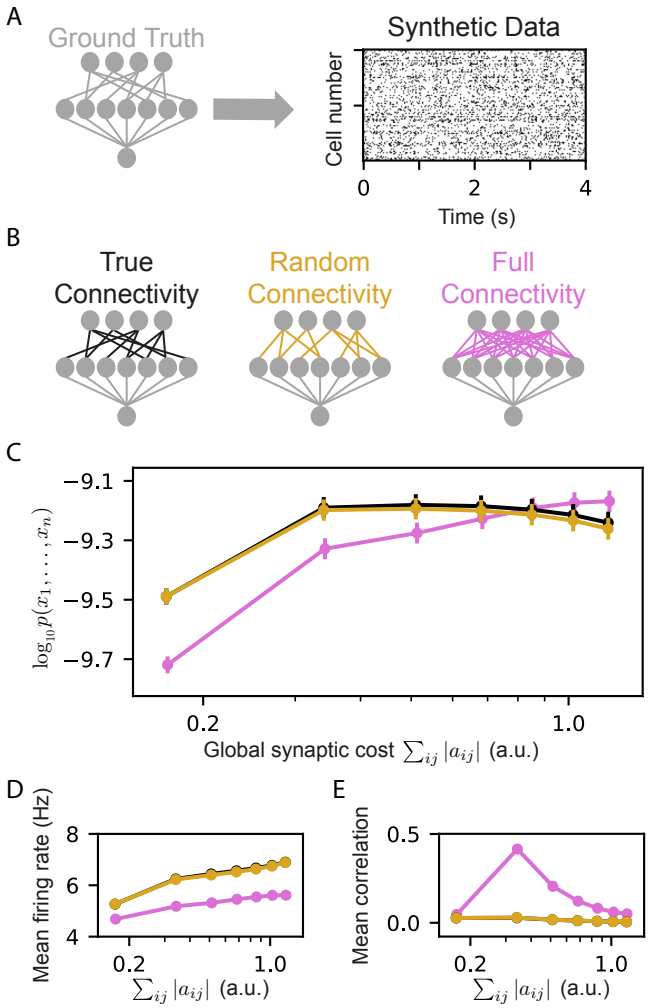


Fig. 4. Models that rely on projections that use random connectivity show similar performance to models that use the correct connectivity. (A) Synthetic population activity data is sampled from an RP model with known connectivity (i.e., the "ground truth" model; see Methods). (B) Homeostatic reshaped random projections models that differ in their connectivity are learned to fit the synthetic data. The "True connectivity" model uses projections whose connectivity is identical to the "ground truth" model. The "Random connectivity" model uses projections that are randomly sampled using sparse random connectivity. The "Full connectivity" model is a homeostatic reshaped model that uses projections with full connectivity. (C) The mean log-likelihood of the models is shown as a function of the model's cost. The true connectivity model is only slightly better than the random connectivity model, with both outperforming the full connectivity model for low model budget values. (D) The mean correlation between the activity of the projection neurons, shown as a function of the model cost. We note that true and random connectivity models are indistinguishable. (E) The firing rates of the projection neurons, shown as a function of the model cost. Note that in panels D, and E, the standard errors are smaller than the markers' size and are, therefore, invisible.

therefore, asked how the initial random "wiring" of the projections affects the performance of the model, and whether non-random projections would result in even better models. To quantify the effects of the projections' connectivity on the performance and efficiency of reshaped models, we used simulated population activity that we generated using RP models that were trained on real data. By using synthetic data that was generated by a known model, we can compare the learned models to the "ground truth" in terms of connectivity, as well as extensively sampling of activity patterns from the model.

We learned homeostatic reshaped models for the synthetic data, using different initial connectivity structures (Figure 4A-B): (i) A "true" connectivity model in which we reshaped a random projections model that has the same connectivity as the projections of the model that generated the data. (ii) A Random connectivity model in which we reshaped projections with sparse and connectivity that is randomly sampled and is independent of the model that generated the synthetic data. (iii) A full connectivity model in which we reshaped random projections with full connectivity, i.e., all input neurons are connected to all the projections, but with random initial weights. We carried out homeostatic reshaping of the projections in all three models with different values of θ. Surprisingly, the true and random connectivity models performed very similarly (Figure 4C). Although the full connectivity model contains the "ground truth" connectivity, and could recreate the true connectivity by canceling out unnecessary synapses during reshaping – we find that the full connectivity models are inferior to the other models, except for the case of high model costs.

The mean correlations between projections at the end of reshaping and the mean firing rates of the models that use the true and random connectivity were also very similar (Figure 4D-E), whereas the full connectivity models showed, again, very different behavior. These results reflect another computational benefit of homeostatic reshaping: there is no need to know the optimal circuit connectivity, and there is no apparent benefit to all-to-all connectivity, which would be expensive in terms of the energetic cost, the space needed, and the biological construction. Thus, starting from random connectivity and optimizing the circuit under homeostatic constraints seems to provide optimal results.

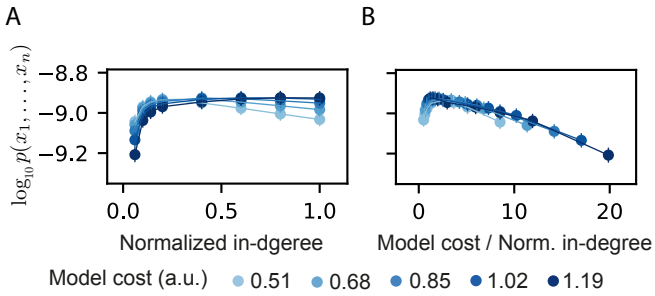


Fig. 5. Fixed ratio between optimal in-degree values of the Reshaped models and their cost. (A) The performance of the tested models, shown as a function of the normalized in-degree of the projections (0–disconnected, 1–fully connected) for different values of the model's cost. (B) The performance of the tested models, shown as a function of the cost, normalized by the in-degree of the projections. Curves of different cost values coincide, suggesting a fixed optimal cost/activity ratio. The standard errors over 100 models are smaller than the size of markers and are, therefore, invisible.

Optimal sparseness of reshaped Projections models under homeostatic constraints . Given the inefficiency of the fully connected reshaped projections model, we quantified the effect of the sparseness of the projections on reshaped RP models. We recall that for the standard RP model, sparse projections were optimal for a wide range of network sizes (30)), and so we measured the performance of homeostatic

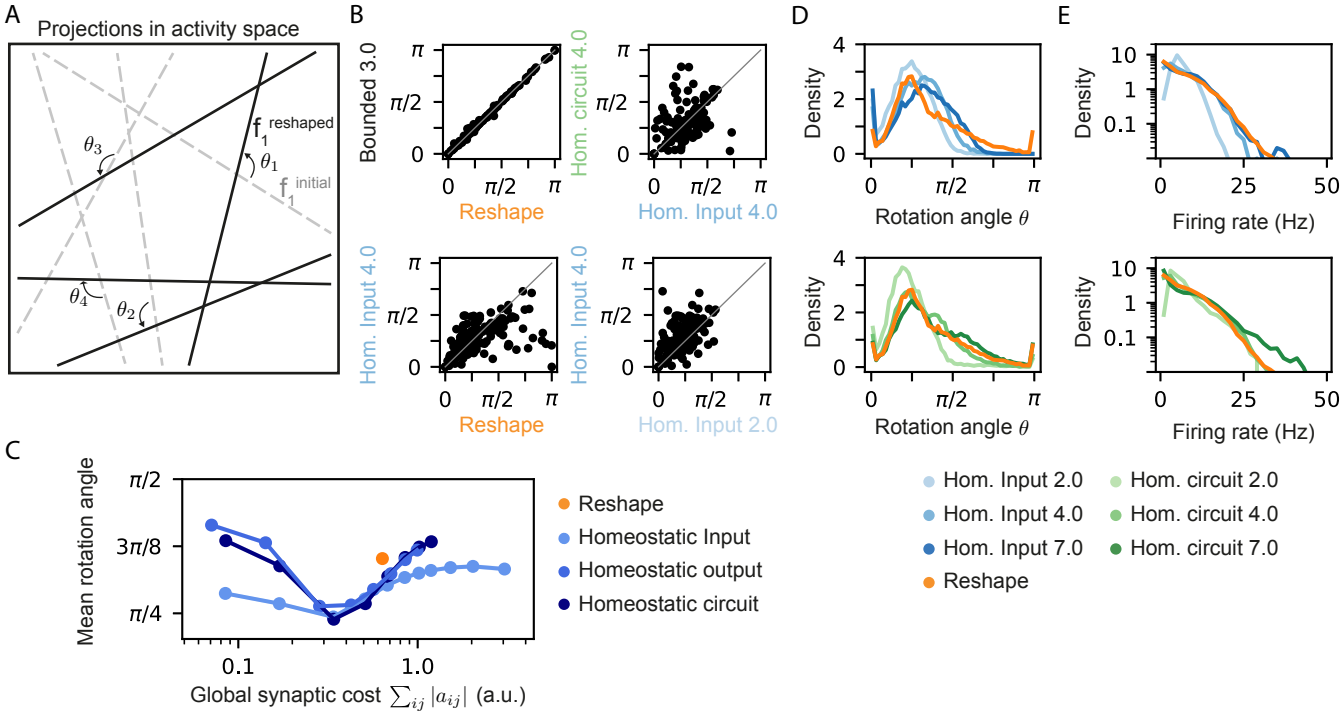


Fig. 6. Different homeostatic models produce a similar behavior using different projection sets. (A) Schematic drawing of how projections rotate during reshaping: starting from the initial projections (grey lines), they rotate to their reshaped orientation (black lines) by angle θ_i. (B) Rotation angles after reshaping, shown for different pairs of models. All four panels show models that initialized with the same set of projections. The different labels specify the constraint type and strength, namely, the specific value of ϕ and θ. (C) The mean rotation angle of the projections due to reshaping, shown as a function of the model synaptic budget. (D) A histogram of rotation angles of the projections after learning in the case of the unconstrained reshape models and different homeostatic input (top) and homeostatic circuit (bottom) models. The legend specifies the constraint type and strength, namely, the specific value of ϕ and θ. (E) A histogram of the firing rates of the same sets of models from (D).

reshaped RP models for different values of in-degree of the projections, while keeping the total synaptic budget of the models fixed. We found that different synaptic budgets have a different optimal in-degree (Figure 5A), and that the value of the optimal in-degree seems to grow with the total synaptic budget.

We further estimated the efficiency of the models by the synaptic cost per connection in the projections (Figure 5B). We find that curves for different total synaptic costs seem to coincide and have a similar peak value – suggesting an optimal ratio between the total available resources and the number of synapses.

Different Homeostatic mechanisms for reshaping random projections models result in different projection sets. Homeostatic reshaping of random projections proved to be highly accurate and efficient for the three homeostatic model variants we tried, and were similar also in their efficiency in terms of total synaptic costs. We, therefore, asked whether we can identify other features that differentiate between the homeostatic models. Since each projection defines a hyperplane in the space of population activity patterns, reshaping can be interpreted as a rotation or a change of the angle of these hyperplanes, depicted schematically in Figure 6A. We, therefore, compared the different homeostatic variants of the reshaped projections models by initializing them from the same set of random projections, and evaluating the corresponding rotation angles, θ, of all of the projections due to the reshaping. Figure 6B shows an example of the rotations of the same initial projections for one model under different reshaping constraints, highlighting the substantial differences between them.

Figure 6C shows the mean rotation angle over 100 homeostatic models as a function of synaptic cost – reflecting that the different forms of homeostatic regulation results in different reshaped projections. Figure 6D shows the histogram of the rotation angles of several homeostatic models, as well as the unconstrained reshape model. Again, we find significant differences between the homeostatic models in terms of the firing rates of the projection neurons after reshaping, as well as for different parameter values of the same homeostatic mechanism (Figure 6E).

Discussion

We presented a new family of statistical models for large neural populations that is based on sparse and random non-linear projections of the population, which are adapted during learning. This new family of models proved to be more accurate than the highly accurate Random Projections class of models, using fewer projections and incurring a lower "synaptic cost" in terms of the total sum of synaptic weights of the model. Moreover, we found that homeostatic reshaping of the projections gave even more accurate and efficient models in terms of synaptic weights, and was optimal for models that relied on random and sparse initial connectivity (compared to fully

connected models).

Our results suggest a computational role for the experimentally observed homeostatic scaling of synapses during learning. Going beyond its customary justification as a "regularizer" of firing rates in neural circuits – in our Reshaped RP models, homeostatic plasticity optimizes the performance of network models and their efficiency in scenarios of limited resources and random connectivity. Moreover, the similarity of the performance of models that use different homeostatic synaptic mechanisms suggests a possible universal role for homeostatic mechanisms in computation.

The shallowness of the circuit implementation of the Reshaped Random Projections model implies that the learning of these models does not require the backpropagation of information over many layers, which distinguishes deep artificial networks from biological ones. Moreover, the locality of the reshaping process itself points to the feasibility of this model in terms of real biological circuits. The biological plausibility is further supported by the robustness of the model to the specific connectivity used for the reshaped models, and to the specific choice of the homeostatic mechanism we used.

A key remaining issue for the biological feasibility of the RP family of models is the feedback signal from the readout neuron to the intermediate neurons. The noise-dependent learning mechanism for RP models presented in (30) and for other local feedback and synaptic learning mechanisms that approximate backprogapation (35) offers clear directions for future study. Our results may also be relevant for learning in artificial neural networks, whose training relies on non-convex approaches that necessitate different regularization techniques (60). The homeostatic mechanism we focused on here is a form of "hard" L1 regularization, but on the sum of the weights. This approach limits the search space, compared to regularization over the weights themselves, but defines coupled changes in weights, in a manner highly effective for the cortical data we studied. We, therefore, hypothesize that homeostatic scaling may be beneficial for artificial architectures (see, e.g., (37)).

Materials and Methods

Experimental Data. Extra-cellular recordings were performed using Utah arrays from populations of neurons in the prefrontal cortex of macaque monkeys performing a direction discrimination task with random dots. For more details see (57).

Data Pre-processing. Neural activity was discretized using 20 ms bins, such that in each time bin a neuron was active ('1') if it emitted a spike in that bin and silent ('0') if not. Recorded data was split randomly into training sets and held-out test sets: 100 different random splits were generated for each model setup, consisting of 160,000 samples in the training set and 40,000 in the test set.

Constructing Sparse Random Projections. Following (30), the coefficients a_{ij} of the random projections are set using a two stage process. First, the connectivity of the projections is set such that the average in-degree of the projections matches a predetermined sparsity value: each input neuron connects to each projection with a probability $p = indegree/n$, where n is the number of neurons in the input layer. The corresponding a_{ij} coefficients are then sampled from a Gaussian distribution, $a_{ij} \sim \mathcal{N}(1,1)$, and the remaining a_{ij} values are set to zero. The threshold of each projection, θ_i, was set to 1.

The average in-degree of sparse models used here was 5, unless specified otherwise in the text. For the fully connected models $indegree = n$ (i.e., sparsity=0).

Training RP models. Given empirical data $\mathbf{X}$ and a set of projections defined by a_{ij}, we train the RP models by searching for the parameters λ_i that maximize the log-likelihood of the model given the data, $\arg\max_{\lambda_i}(L(\mathbf{X}))$, where $L(\mathbf{X}) = \sum_{\vec{x}\in\mathbf{X}} \log p_{RP}(\vec{x})$. This is a convex function whose gradient is given by

$$\nabla_{\lambda_i} L(\mathbf{X}) = \langle f_i \rangle_{\mathbf{X}} - \langle f_i \rangle_{p_{RP}} . \quad \textbf{(6)}$$

We found the values λ_i that maximize the log-likelihood by gradient descent with momentum or ADAM algorithms. We computed the empirical expectation in $\langle f_i \rangle_{\mathbf{X}}$ by summing over the training data, and the expectation over the probability model $\langle f_i \rangle_{p_{RP}}$ by summing over synthetic data generated from p_{RP} using Metropolis–Hasting sampling.

For each of the empirical marginals $\langle f_i \rangle_{\mathbf{X}}$, we used the Clopper–Pearson method to estimate the distribution of possible values for the real marginal given the empirical observation. We set the convergence threshold of the numerical solver such that each of the marginals in the model distribution falls within a CI of one SD under this distribution, from its empirical marginal.

Reshaping RP models. Given empirical data $\mathbf{X}$, we optimize the RP models by modifying the coefficients a_{ij} such that the log-likelihood of the model is maximized, $\arg\max_{a_{ij}}(L(\mathbf{X}))$. Starting from an initial set of projections, a^0_{ij}, using the update rule of equation 3, we optimize the projections by applying the gradient descent with momentum algorithm. Importantly, only non-zero elements of a^0_{ij} are optimized.

Optimizing backpropagation models. Full backpropagation models are optimized using the learning rules of the trained RP models and the reshaped models simultaneously in each gradient descent step, i.e., eqs. 3 and 6.

Homeostatic reshaping of RP models. The homeostatic RP models are reshaped as follows: We first define a set of unconstrained projections where the coefficients $\tilde{a}_{ij}$ are randomly sampled. Each of the projections is then normalized homeostatically, such that a_{ij} are a function of this unconstrained set: $a_{ij} = \phi \cdot \tilde{a}_{ij} / \sum_k |\tilde{a}_{ik}|$, where ϕ is the available synaptic budget for each projection. We then optimize $\tilde{a}_{ij}$ to maximize the log-likelihood of the model given the empirical data $\mathbf{X}$: $\arg\max_{\tilde{a}_{ij}}(L(\mathbf{X}))$. The computed constrained

projections a_{ij} are then used in the resulting homeostatic RP model.

Bounded reshaping of RP models. Similar to reshaping homeostatic RP models, we define a set of unconstrained projections $\tilde{a}_{ij}$, where the projections are a function of this unconstrained set: $a_{ij} = \min\left(\max\left(\tilde{a}_{ij}, -\theta\right), \theta\right)$, where θ is the "ceiling" value of each synapse.

Generating synthetic data from RP models with known connectivity. Synthetic neural activity patterns were obtained by training RP models on real neural recordings as described above and then generating data from these models using Metropolis-Hastings sampling.

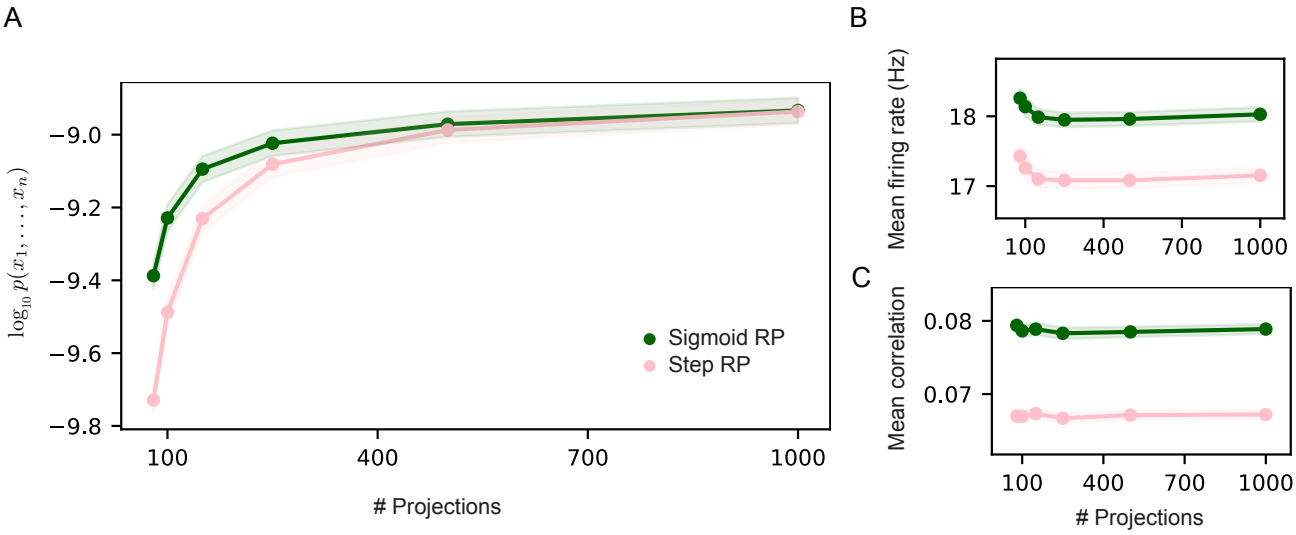
ACKNOWLEDGEMENTS
We thank Adam Haber, Tal Tamir, Udi Karpas, and the rest of the Schneidman lab members for discussions, comments, and ideas. This work was supported by Simons Collaboration on the Global Brain grant 542997 (ES), Israel Science Foundation grant 137628 (ES), Israeli Council for Higher Education/Weizmann Data Science Research Center (ES), Martin Kushner Schnur, and Mr. & Mrs. Lawrence Feis. ES is the incumbent of the Joseph and Bessie Feinberg Chair.

Bibliography

1. Wilson Truccolo, Uri T. Eden, Matthew R. Fellows, John P. Donoghue, and Emery N. Brown. A Point Process Framework for Relating Neural Spiking Activity to Spiking History, Neural Ensemble, and Extrinsic Covariate Effects. *Journal of Neurophysiology*, 93(2):1074–1089, February 2005. ISSN 0022-3077, 1522-1598. doi: 10.1152/jn.00697.2004.
2. Jonathan W. Pillow, Jonathon Shlens, Liam Paninski, Alexander Sher, Alan M. Litke, E. J. Chichilnisky, and Eero P. Simoncelli. Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, 454(7207):995–999, August 2008. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature07140.
3. Ana Calabrese, Joseph W. Schumacher, David M. Schneider, Liam Paninski, and Sarah M. N. Woolley. A Generalized Linear Model for Estimating Spectrotemporal Receptive Fields from Responses to Natural Sounds. *PLoS ONE*, 6(1):e16104, January 2011. ISSN 1932-6203. doi: 10.1371/journal.pone.0016104.
4. Franz Weber, Christian K. Machens, and Alexander Borst. Disentangling the functional consequences of the connectivity between optic-flow processing neurons. *Nature Neuroscience*, 15(3):441–448, March 2012. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn.3044.
5. Elad Schneidman, Michael J. Berry, Ronen Segev, and William Bialek. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature*, 440 (7087):1007–1012, April 2006. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature04701.
6. J. Shlens, G. D. Field, J. L. Gauthier, M. I. Grivich, D. Petrusca, A. Sher, A. M. Litke, and E. J. Chichilnisky. The Structure of Multi-Neuron Firing Patterns in Primate Retina. *Journal of Neuroscience*, 26(32):8254–8266, August 2006. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.1282-06.2006.
7. A. Tang, D. Jackson, J. Hobbs, W. Chen, J. L. Smith, H. Patel, A. Prieto, D. Petrusca, M. I. Grivich, A. Sher, P. Hottowy, W. Dabrowski, A. M. Litke, and J. M. Beggs. A Maximum Entropy Model Applied to Spatial and Temporal Correlations from Cortical Networks In Vitro. *Journal of Neuroscience*, 28(2):505–518, January 2008. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.3359-07.2008.
8. Gašper Tkačik, Olivier Marre, Dario Amodei, Elad Schneidman, William Bialek, and Michael J. Berry. Searching for Collective Behavior in a Large Network of Sensory Neurons. *PLoS Computational Biology*, 10(1):e1003408, January 2014. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1003408.
9. Elad Ganmor, Ronen Segev, and Elad Schneidman. Sparse low-order interaction network underlies a highly correlated and learnable neural population code. *Proceedings of the National Academy of Sciences*, 108(23):9679–9684, June 2011. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1019641108.
10. O. Marre, S. El Boustani, Y. Frégnac, and A. Destexhe. Prediction of Spatiotemporal Patterns of Neural Activity from Pairwise Correlations. *Physical Review Letters*, 102(13): 138101, April 2009. doi: 10.1103/PhysRevLett.102.138101.
11. Ifije E. Ohiorhenuan, Ferenc Mechler, Keith P. Purpura, Anita M. Schmid, Qin Hu, and Jonathan D. Victor. Sparse coding and high-order correlations in fine-scale cortical networks. *Nature*, 466(7306):617–621, July 2010. ISSN 1476-4687. doi: 10.1038/nature09178.
12. Einat Granot-Atedgi, Gašper Tkačik, Ronen Segev, and Elad Schneidman. Stimulus-dependent Maximum Entropy Models of Neural Population Codes. *PLOS Computational Biology*, 9(3):e1002922, March 2013. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1002922.
13. Leenoy Meshulam, Jeffrey L. Gauthier, Carlos D. Brody, David W. Tank, and William Bialek. Collective Behavior of Place and Non-place Neurons in the Hippocampal Network. *Neuron*, 96(5):1178–1191.e4, December 2017. ISSN 0896-6273. doi: 10.1016/j.neuron.2017.10.027.
14. Elad Ganmor, Ronen Segev, and Elad Schneidman. A thesaurus for a neural population code. *eLife*, 4:e06134, September 2015. ISSN 2050-084X. doi: 10.7554/eLife.06134.
15. Gašper Tkačik, Einat Granot-Atedgi, Ronen Segev, and Elad Schneidman. Retinal Metric: A Stimulus Distance Measure Derived from Population Neural Responses. *Physical Review Letters*, 110(5):058104, January 2013. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.110.058104.
16. Chethan Pandarinath, Daniel J. O'Shea, Jasmine Collins, Rafal Jozefowicz, Sergey D. Stavisky, Jonathan C. Kao, Eric M. Trautmann, Matthew T. Kaufman, Stephen I. Ryu, Leigh R. Hochberg, Jaimie M. Henderson, Krishna V. Shenoy, L. F. Abbott, and David Sussillo. Inferring single-trial neural population dynamics using sequential auto-encoders. *Nature Methods*, 15(10):805–815, October 2018. ISSN 1548-7105. doi: 10.1038/s41592-018-0109-9.
17. David GT Barrett, Ari S Morcos, and Jakob H Macke. Analyzing biological and artificial neural networks: challenges with opportunities for synergy? *Current Opinion in Neurobiology*, 55:55–64, April 2019. ISSN 0959-4388. doi: 10.1016/j.conb.2019.01.007.
18. Pedro J Gonçalves, Jan-Matthis Lueckmann, Michael Deistler, Marcel Nonnenmacher, Kaan Öcal, Giacomo Bassetto, Chaitanya Chintaluri, William F Podlaski, Sara A Haddad, Tim P Vogels, David S Greenberg, and Jakob H Macke. Training deep neural density estimators to identify mechanistic models of neural dynamics. *eLife*, 9:e56261, September 2020. ISSN 2050-084X. doi: 10.7554/eLife.56261.
19. Gašper Tkačik, Thierry Mora, Olivier Marre, Dario Amodei, Stephanie E. Palmer, Michael J. Berry, and William Bialek. Thermodynamics and signatures of criticality in a network of neurons. *Proceedings of the National Academy of Sciences*, 112(37):11508–11513, September 2015. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1514188112.
20. Leenoy Meshulam, Jeffrey L. Gauthier, Carlos D. Brody, David W. Tank, and William Bialek. Coarse Graining, Fixed Points, and Scaling in a Large Population of Neurons. *Physical Review Letters*, 123(17):178103, October 2019. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.123.178103.
21. Elad Schneidman. Towards the design principles of neural population codes. *Current Opinion in Neurobiology*, 37:133–140, April 2016. ISSN 0959-4388. doi: 10.1016/j.conb.2016.03.001.
22. Ed Karpas, O Maoz, R Kiani, and E Schneidman. Strongly correlated spatiotemporal encoding and simple decoding in the prefrontal cortex. preprint, Neuroscience, July 2019.
23. Stefano Panzeri, Christopher D. Harvey, Eugenio Piasini, Peter E. Latham, and Tommaso Fellin. Cracking the Neural Code for Sensory Perception by Combining Statistics, Intervention, and Behavior. *Neuron*, 93(3):491–507, February 2017. ISSN 08966273. doi: 10.1016/j.neuron.2016.12.036.
24. Gašper Tkačik, Olivier Marre, Thierry Mora, Dario Amodei, Michael J. Berry II, and William Bialek. The simplest maximum entropy model for collective behavior in a neural network. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(03):P03011, March 2013. ISSN 1742-5468. doi: 10.1088/1742-5468/2013/03/P03011.
25. Vicente Botella-Soler, Stéphane Deny, Georg Martius, Olivier Marre, and Gašper Tkačik. Nonlinear decoding of a complex movie from the mammalian retina. *PLOS Computational Biology*, 14(5):e1006057, May 2018. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1006057.
26. Qing Shi, Pranjal Gupta, Alexandra K. Boukhvalova, Joshua H. Singer, and Daniel A. Butts. Functional characterization of retinal ganglion cells using tailored nonlinear modeling. *Scientific Reports*, 9(1):8713, June 2019. ISSN 2045-2322. doi: 10.1038/s41598-019-45048-8.
27. Matthew R. Whiteway, Bruno Averbeck, and Daniel A. Butts. A latent variable approach to decoding neural population activity, January 2020. Pages: 2020.01.06.896423 Section: New Results.
28. Juan A. Gallego, Matthew G. Perich, Raeed H. Chowdhury, Sara A. Solla, and Lee E. Miller. Long-term stability of cortical population dynamics underlying consistent behavior. *Nature Neuroscience*, 23(2):260–270, February 2020. ISSN 1546-1726. doi: 10.1038/s41593-019-0555-4.
29. Rishidev Chaudhuri, Berk Gerçek, Biraj Pandey, Adrien Peyrache, and Ila Fiete. The intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleep. *Nature Neuroscience*, 22(9):1512–1520, September 2019. ISSN 1546-1726. doi: 10.1038/s41593-019-0460-x.
30. Ori Maoz, Gašper Tkačik, Mohamad Saleh Esteki, Roozbeh Kiani, and Elad Schneidman. Learning probabilistic neural representations with randomly connected circuits. *Proceedings of the National Academy of Sciences*, 117(40):25066–25073, October 2020. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1912804117.
31. Eszter Vertes and Maneesh Sahani. Flexible and accurate inference and learning for deep generative models, May 2018. arXiv:1805.11051 [cs, stat].
32. Richard S. Zemel, Peter Dayan, and Alexandre Pouget. Probabilistic Interpretation of Population Codes. *Neural Computation*, 10(2):403–430, February 1998. ISSN 0899-7667, 1530-888X. doi: 10.1162/089976698300017818.
33. Yoshua Bengio, Dong-Hyun Lee, Jorg Bornschein, Thomas Mesnard, and Zhouhan Lin. Towards Biologically Plausible Deep Learning, August 2016. arXiv:1502.04156 [cs].
34. Daniel L K Yamins and James J DiCarlo. Using goal-driven deep learning models to understand sensory cortex. *Nature Neuroscience*, 19(3):356–365, March 2016. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn.4244.
35. Panayiota Poirazi, Terrence Brannon, and Bartlett W. Mel. Pyramidal Neuron as Two-Layer Neural Network. *Neuron*, 37(6):989–999, March 2003. ISSN 0896-6273. doi: 10.1016/S0896-6273(03)00149-1.
36. Blake A. Richards, Timothy P. Lillicrap, Philippe Beaudoin, Yoshua Bengio, Rafal Bogacz, Amelia Christensen, Claudia Clopath, Rui Ponte Costa, Archy de Berker, Surya Ganguli, Colleen J. Gillon, Danijar Hafner, Adam Kepecs, Nikolaus Kriegeskorte, Peter Latham, Grace W. Lindsay, Kenneth D. Miller, Richard Naud, Christopher C. Pack, Panayiota Poirazi, Pieter Roelfsema, João Sacramento, Andrew Saxe, Benjamin Scellier, Anna C. Schapiro, Walter Senn, Greg Wayne, Daniel Yamins, Friedemann Zenke, Joel Zylberberg, Denis Therien, and Konrad P. Kording. A deep learning framework for neuroscience. *Nature Neuroscience*, 22(11):1761–1770, November 2019. ISSN 1546-1726. doi: 10.1038/s41593-019-0520-2.
37. Weishun Zhong, Ben Sorscher, Daniel Lee, and Haim Sompolinsky. A theory of weight distribution-constrained learning. In Alice H. Oh, Alekh Agarwal, Danielle Belgrave, and Kyunghyun Cho, editors, *Advances in Neural Information Processing Systems*, 2022.
38. Spyridon Chavlis and Panayiota Poirazi. Drawing inspiration from biological dendrites to empower artificial neural networks. *Current Opinion in Neurobiology*, 70:1–10, October 2021. ISSN 0959-4388. doi: 10.1016/j.conb.2021.04.007.
39. Ashok Litwin-Kumar, Kameron Decker Harris, Richard Axel, Haim Sompolinsky, and L. F. Abbott. Optimal Degrees of Synaptic Connectivity. *Neuron*, 93(5):1153–1164.e7, 2017.

ISSN 0896-6273. doi: https://doi.org/10.1016/j.neuron.2017.01.030.
40. Adam Haber and Elad Schneidman. Learning the Architectural Features That Predict Functional Similarity of Neural Networks. *Physical Review X*, 12(2):021051, June 2022. ISSN 2160-3308. doi: 10.1103/PhysRevX.12.021051.
41. Sung Soo Kim, Ann M. Hermundstad, Sandro Romani, L. F. Abbott, and Vivek Jayaraman. Generation of stable heading representations in diverse visual scenes. *Nature*, 576(7785): 126–131, December 2019. ISSN 1476-4687. doi: 10.1038/s41586-019-1767-1.
42. Vladyslava Pechuk, Gal Goldman, Yehuda Salzberg, Aditi H. Chaubey, R. Aaron Bola, Jonathon R. Hoffman, Morgan L. Endreson, Renee M. Miller, Noah J. Reger, Douglas S. Portman, Denise M. Ferkey, Elad Schneidman, and Meital Oren-Suissa. Reprogramming the topology of the nociceptive circuit in C. elegans reshapes sexual behavior. *Current Biology*, 32(20):4372–4385.e7, October 2022. ISSN 09609822. doi: 10.1016/j.cub.2022.08.038.
43. Adam Haber and Elad Schneidman. The computational and learning benefits of Daleian neural networks, October 2022. arXiv:2210.05961 [q-bio].
44. David J. Heeger. Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, 9(2):181–197, August 1992. ISSN 0952-5238, 1469-8714. doi: 10.1017/S0952523800009640.
45. Matteo Carandini and David J. Heeger. Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, 13(1):51–62, January 2012. ISSN 1471-003X, 1471-0048. doi: 10.1038/nrn3136.
46. Gina G. Turrigiano, Kenneth R. Leslie, Niraj S. Desai, Lana C. Rutherford, and Sacha B. Nelson. Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature*, 391(6670):892–896, February 1998. ISSN 1476-4687. doi: 10.1038/36103.
47. Tara Keck, Georg B. Keller, R. Irene Jacobsen, Ulf T. Eysel, Tobias Bonhoeffer, and Mark Hübener. Synaptic Scaling and Homeostatic Plasticity in the Mouse Visual Cortex In Vivo. *Neuron*, 80(2):327–334, October 2013. ISSN 0896-6273. doi: 10.1016/j.neuron.2013.08.018.
48. Keith B. Hengen, Mary E. Lambo, Stephen D. Van Hooser, Donald B. Katz, and Gina G. Turrigiano. Firing Rate Homeostasis in Visual Cortex of Freely Behaving Rodents. *Neuron*, 80(2):335–342, October 2013. ISSN 0896-6273. doi: 10.1016/j.neuron.2013.08.038.
49. Gina G. Turrigiano. The Self-Tuning Neuron: Synaptic Scaling of Excitatory Synapses. *Cell*, 135(3):422–435, October 2008. ISSN 00928674. doi: 10.1016/j.cell.2008.10.008.
50. Sami El-Boustani, Jacque P. K. Ip, Vincent Breton-Provencher, Graham W. Knott, Hiroyuki Okuno, Haruhiko Bito, and Mriganka Sur. Locally coordinated synaptic plasticity of visual cortex neurons in vivo. *Science*, 360(6395):1349–1354, June 2018. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.aao0862.
51. Yue Kris Wu, Keith B. Hengen, Gina G. Turrigiano, and Julijana Gjorgjieva. Homeostatic mechanisms regulate distinct aspects of cortical circuit dynamics. *Proceedings of the National Academy of Sciences*, 117(39):24514–24525, September 2020. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1918368117.
52. Tara Keck, Taro Toyoizumi, Lu Chen, Brent Doiron, Daniel E. Feldman, Kevin Fox, Wulfram Gerstner, Philip G. Haydon, Mark Hübener, Hey-Kyoung Lee, John E. Lisman, Tobias Rose, Frank Sengpiel, David Stellwagen, Michael P. Stryker, Gina G. Turrigiano, and Mark C. van Rossum. Integrating Hebbian and homeostatic plasticity: the current state of the field and future research directions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1715):20160158, March 2017. ISSN 0962-8436, 1471-2970. doi: 10.1098/rstb.2016.0158.
53. Friedemann Zenke and Wulfram Gerstner. Hebbian plasticity requires compensatory processes on multiple timescales. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1715):20160259, March 2017. ISSN 0962-8436, 1471-2970. doi: 10.1098/rstb.2016.0259.
54. Taro Toyoizumi, Megumi Kaneko, Michael P. Stryker, and Kenneth D. Miller. Modeling the Dynamic Interaction of Hebbian and Homeostatic Plasticity. *Neuron*, 84(2):497–510, October 2014. ISSN 0896-6273. doi: 10.1016/j.neuron.2014.09.036.
55. Eero P. Simoncelli and David J. Heeger. A model of neuronal responses in visual area MT. *Vision Research*, 38(5):743–761, 1998. ISSN 0042-6989. doi: https://doi.org/10.1016/S0042-6989(97)00183-1.
56. E. T. Jaynes. Information Theory and Statistical Mechanics. *Physical Review*, 106(4):620–630, May 1957. ISSN 0031-899X. doi: 10.1103/PhysRev.106.620.
57. Roozbeh Kiani, Christopher J. Cueva, John B. Reppas, and William T. Newsome. Dynamics of Neural Population Responses in Prefrontal Cortex Indicate Changes of Mind on Single Trials. *Current Biology*, 24(13):1542–1547, July 2014. ISSN 09609822. doi: 10.1016/j.cub.2014.05.049.
58. Horace B Barlow. Possible principles underlying the transformation of sensory messages. *Sensory communication*, 1(01), 1961.
59. Bruno A. Olshausen and David J. Field. Sparse coding with an overcomplete basis set: A strategy employed by V1? *Vision Research*, 37(23):3311–3325, December 1997. ISSN 00426989. doi: 10.1016/S0042-6989(97)00169-7.
60. Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.

Supplementary Note 1: Step and Sigmoid Random Projections models comparison



Supp. Fig. 1. Sigmoid Random Projections models outperform Step Random Projections models. (A) Average performance of models as a function of the number of projections, measured by log-likelihood. Sigmoid RP models outperform step RP models. (B-C) Mean firing rates of projection neurons and mean correlation between projections, step RP models show slightly lower correlations and firing rates compared to sigmoid RP models. In all panels shaded area denotes standard error over 100 models.