Shaping Dynamics With Multiple Populations in Low-Rank Recurrent Networks

Manuel Beiran
manuel.beiran@ens.fr
Alexis Dubreuil
alexis.dubreuil@gmail.com
Adrian Valente
adrian.valente@ens.fr
Laboratoire de Neurosciences Cognitives et Computationnelles, INSERM U960,
Ecole Normale Superieure. PSL University, 75005 Paris, France

Francesca Mastrogiuseppe
f.mastrogiuseppe@ucl.ac.uk
Gatsby Computational Neuroscience Unit, UCL, London W1T 4JG, U.K.

Srdjan Ostojic
srdjan.ostojic@ens.fr
Laboratoire de Neurosciences Cognitives et Computationnelles, INSERM U960,
Ecole Normale Superieure. PSL University, 75005 Paris, France

An emerging paradigm proposes that neural computations can be understood at the level of dynamic systems that govern low-dimensional trajectories of collective neural activity. How the connectivity structure of a network determines the emergent dynamical system, however, remains to be clarified. Here we consider a novel class of models, gaussian-mixture, low-rank recurrent networks in which the rank of the connectivity matrix and the number of statistically defined populations are independent hyperparameters. We show that the resulting collective dynamics form a dynamical system, where the rank sets the dimensionality and the population structure shapes the dynamics. In particular, the collective dynamics can be described in terms of a simplified effective circuit of interacting latent variables. While having a single global population strongly restricts the possible dynamics, we demonstrate that if the number of populations is large enough, a rank R network can approximate any R-dimensional dynamical system.

1 Introduction

A newly emerging paradigm posits that neural computations rely on collective dynamics in the state-space corresponding to the joint activity of all

neurons in a network (Churchland & Shenoy, 2007; Rabinovich, Huerta, & Laurent, 2008; Buonomano & Maass, 2009; Saxena & Cunningham, 2019; Vyas, Golub, Sussillo, & Shenoy, 2020). Experiments in behaving animals have found that trajectories of neural activity are typically restricted to low-dimensional manifolds in that space (Machens, Romo, & Brody, 2010; Mante, Sussillo, Shenoy, & Newsome, 2013; Rigotti et al., 2013; Gao, Ganguli, Battaglia, & Schnitzer, 2015; Gallego et al., 2018; Chaisangmongkon, Swaminathan, Freedman, & Wang, 2017; Wang, Narain, Hosseini, & Jazayeri, 2018; Sohn, Narain, Meirhaeghe, & Jazayeri, 2019) and can therefore be described by a small number of collective, latent variables. It has been proposed that these collective variables form dynamical systems that implement computations through their responses to inputs (Eliasmith & Anderson, 2003; Hennequin, Vogels, & Gerstner, 2014; Rajan, Harvey, & Tank, 2016; Remington, Egger, Narain, Wang, & Jazayeri, 2018; Remington, Narain, Hosseini, & Jazayeri, 2018). How synaptic connectivity shapes the effective dynamics of collective variables, and therefore computations, however remains to be clarified.

Recurrent neural networks (RNNs) trained to perform neuroscience tasks are an ideal model system to address this question and further develop the theory of computations through dynamics (Sussillo, Churchland, Kaufman, & Shenoy, 2015; Rajan et al., 2016; Barak, 2017; Wang et al., 2018; Yang, Joglekar, Song, Newsome, & Wang, 2019). A recently introduced class of models, low-rank RNNs, directly embodies the idea of low-dimensional collective dynamics, opens the door to relating connectivity and dynamics, and provides a framework that unifies a number of specific RNN classes (Mastrogiuseppe & Ostojic, 2018). Low-rank RNNs rely on connectivity matrices that are restricted to be low rank, which directly generate low-dimensional dynamics. The rank of the network determines the number of collective variables needed to provide a full description of the collective dynamics. While previous works have shown that other specific classes of RNNs can approximate arbitrary dynamical systems (Doya, 1993; Maass, Joshi, & Sontag, 2007), the range of collective dynamics that can be implemented by low-rank RNNs remains to be clarified.

In this work, we focus on low-rank RNNs in which neurons are organized in distinct populations that correspond to clusters in the space of low-rank connectivity patterns. Each population is defined by its statistics of connectivity, described by a multivariate gaussian distribution, so that the full network is specified by a mixture of gaussians. The total number of populations in the network is a hyperparameter distinct from the rank of connectivity. Previous works have considered low-rank networks consisting of a single, global gaussian population (Mastrogiuseppe & Ostojic, 2018, 2019; Schuessler, Dubreuil, Mastrogiuseppe, Ostojic, & Barak, 2020). In the opposite limit, by increasing the number of populations, a gaussian mixture model can approximate any arbitrary low-rank connectivity distribution. Here we examine how the number of populations and

their structure determine and limit the resulting collective dynamics in the network.

We first derive three general properties of gaussian-mixture low-rank networks: (1) in an autonomous network of rank R, dynamics are characterized by R collective variables that form a dynamical system; (2) the dynamics are determined by an effective circuit description, where collective variables interact through gain-modulated effective couplings; and (3) the resulting low-dimensional dynamics can approximate any arbitrary R-dimensional dynamical system if the number of populations is large enough. We then illustrate how increasing the number of populations in a network extends its dynamical range. For that, we specifically focus on fixed points of the dynamics. While a network consisting of a single population can generate at most a pair of stable fixed points, independently of its rank, we show that adding populations allows the network to implement arbitrary numbers of stable fixed points embedded in a subspace determined by the rank of the connectivity matrix. Finally, we propose a general algorithm to approximate a given R-dimensional dynamical system with a multipopulation network of rank R and show one example network that is designed to implement complex temporal dynamics.

2 Model Class: Gaussian Mixture Low-Rank Networks

In this section, we introduce the class of models we study and define the key underlying quantities.

We consider a recurrent neural network of N rate units. The dynamics of the input x_i to the ith unit are given by

$$\tau \frac{dx_i}{dt} = -x_i + \sum_{j=1}^{N} J_{ij}\phi(x_i) + I_i^{ext}(t),$$

(2.1)

where τ corresponds to the membrane time constant, the matrix element J_{ij} is the synaptic strength from unit j to unit i, and $I_i^{ext}(t)$ is the external input received by the ith unit. The nonlinear function $\phi(x)$ maps the input of a neuron to its firing rate activity. Throughout this study, we use the nonlinear activation function $\phi(x) = \tanh(x)$, although the theoretical results in section 3 hold for any nonpolynomial activation function.

We restrict the connectivity matrix to be of low rank, that is, the number of nonzero singular values of the matrix J is $R \ll N$. Using singular value decomposition, any connectivity matrix of this type can be expressed as the sum of R unit rank terms:

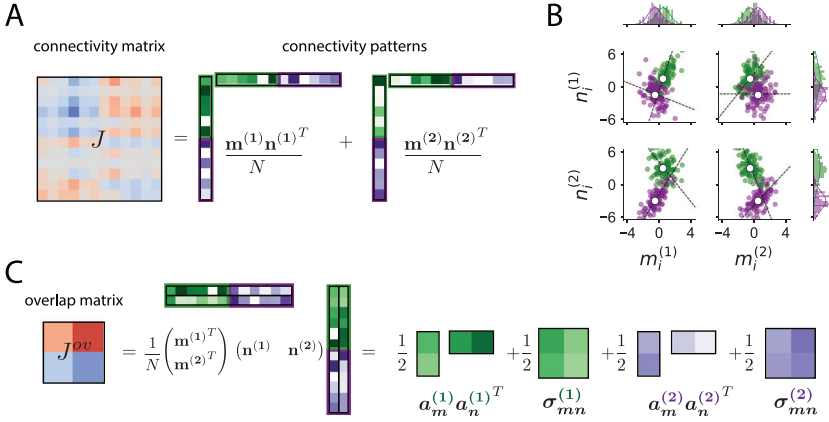$$J_{ij} = \frac{1}{N} \sum_{r=1}^{R} m_i^{(r)} n_j^{(r)}.$$

(2.2)

Figure 1: Low-rank connectivity with gaussian populations. (A) The connectivity matrix J, rank-two in this illustration, is decomposed into the sum of two rank-one terms given by the outer product of the connectivity patterns $\mathbf{m}^{(r)}$ and $\mathbf{n}^{(r)}$, $r = 1, 2$. The components of the connectivity patterns—the pattern loadings—are grouped into two different subpatterns (green and purple) with different population statistics. For visual purposes, the connectivity is shown only for 12 neurons in each population; the first 12 neurons belong to population 1, and the last 12 neurons belong to population 2. (B) Scatter plot of the distribution of pattern components in the four-dimensional loading space. Each dot corresponds to one neuron, and each neuron is characterized by its four values on the patterns $\mathbf{m}^{(r)}$ and $\mathbf{n}^{(r)}$, $r = 1, 2$. The color indicates whether the neuron belongs to the first population (green) or the second population (purple). The different populations are defined by different multivariate gaussian statistics, means (white dots) and covariances (dashed lines), and define separate clusters. Population size $N = 200$, $\alpha_p = 0.5$. (C) Overlap matrix given by the inner product between connectivity patterns. The overlap matrix is a square matrix of size given by the rank of the connectivity, in this case 2×2. Its eigenvalues coincide with the nonzero eigenvalues of the $N \times N$ connectivity matrix. The overlap matrix can be expressed as a weighted sum over the overlaps of the different populations, as shown in equation 2.12.

The connectivity is therefore characterized by a set of R N-dimensional vectors, or connectivity patterns, $\mathbf{m}^{(r)} = \left\{m_i^{(r)}\right\}_{i=1...N}$ and $\mathbf{n}^{(r)} = \left\{n_i^{(r)}\right\}_{i=1...N}$ for $r = 1, \ldots, R$, where $\mathbf{m}^{(r)}$ are the left singular vectors of the connectivity matrix and $\mathbf{n}^{(r)}$ correspond to the right singular vectors multiplied by the corresponding singular values (see Figure 1A for an example of a rank-two connectivity matrix). The vectors $\mathbf{m}^{(r)}$ (resp. $\mathbf{n}^{(r)}$) for $r = 1, \ldots, R$ are mutually orthogonal. Without loss of generality, we fix the norm of the left

singular vectors $\mathbf{m}^{(\mathbf{r})}$ to be equal to N. This decomposition is unique, up to a change in sign of the set of vectors $\mathbf{m}^{(\mathbf{r})}$ and $\mathbf{n}^{(\mathbf{r})}$.

The external input can be expressed as the sum of N_{in} time-varying terms,

$$I_i^{ext}(t) = \sum_{s=1}^{N_{in}} I_i^{(s)} u_s(t), \tag{2.3}$$

which are fed into the network through a set of orthonormal input patterns $\mathbf{I}^{(\mathbf{s})} = \left\{ I_i^{(s)} \right\}_{i=1...N}$ for $s = 1, \ldots, N_{in}$. In this study, we focus on the dynamics of autonomous networks or networks with a constant external input.

Each neuron in the network is therefore characterized by its $2R + N_{in}$ components on the connectivity patterns $\mathbf{m}^{(\mathbf{r})}$ and $\mathbf{n}^{(\mathbf{r})}$ and input patterns $\mathbf{I}^{(\mathbf{s})}$. By analogy with factor analysis, we refer to these components as pattern loadings and denote the set of loadings for neuron i as

$$\left(\left\{ m_i^{(r)} \right\}_{r=1...R}, \left\{ n_i^{(r)} \right\}_{r=1...R}, \left\{ I_i^{(s)} \right\}_{s=1...N_{in}} \right) := (\underline{m}_i, \underline{n}_i, \underline{I}_i). \tag{2.4}$$

Each neuron can thus be represented as a point in the loading space of dimension $2R + N_{in}$, and the connectivity of the full network can therefore be described as a set of N points in this pattern loading space (see Figure 1B).

We assume that for each neuron, the set of pattern loadings is generated independently from a multivariate probability distribution $P(\underline{m}, \underline{n}, \underline{I})$. We moreover restrict ourselves to a specific class of loading distributions: mixtures of multivariate gaussians. This choice is motivated by the fact that gaussian mixtures can approximate any arbitrary multivariate distribution, afford a natural interpretation in terms of populations, and allow for a mathematically tractable and transparent analysis of the dynamics as shown below.

In this gaussian mixture model, each neuron is assigned to a population p with probability α_p, $p = 1 \ldots P$, so that the connectivity matrix J is a block matrix. Within population p, the joint distribution $P^{(p)}(\underline{m}, \underline{n}, \underline{I})$ is a multivariate gaussian defined by (1) its mean $\boldsymbol{a}^{(\boldsymbol{p})}$, a vector of dimension $2R + N_{in}$, given by the set of means of each pattern loading within population p,

$$\boldsymbol{a}^{(\boldsymbol{p})} = \left(a_{m_1}^{(p)}, \ldots, a_{m_R}^{(p)}, a_{n_1}^{(p)}, \ldots, a_{n_R}^{(p)}, a_{I_1}^{(p)}, \ldots, a_{I_{N_{in}}}^{(p)} \right), \tag{2.5}$$

and (2) its covariance $\Sigma^{(p)}$, a matrix of dimension $(2R + N_{in}) \times (2R + N_{in})$, whose elements are the pairwise covariances

$$\Sigma_{xy}^{(p)} = E \left[\left(x^{(p)} - a_x^{(p)} \right) \left(y^{(p)} - a_y^{(p)} \right) \right], \tag{2.6}$$

where $E[\cdot]$ indicates the expected value and x and y represent any pair of connectivity or input components. Within the loading space, each population therefore corresponds to a cluster centered at $a^{(p)}$ and of shape specified by the connectivity matrix $\Sigma_{xy}^{(p)}$ (see Figure 1B).

The geometrical arrangement between patterns is a key feature to understand the behavior of low-rank networks (Mastrogiuseppe & Ostojic, 2018). The connectivity and input patterns are N-dimensional vectors. To quantify the geometrical configuration between two patterns, we define the overlap, or normalized scalar product,

$$O(\mathbf{x}, \mathbf{y}) = \frac{1}{N} \sum_{i=1}^{N} x_i y_i, \tag{2.7}$$

where $\mathbf{x}$ and $\mathbf{y}$ are any two patterns in the set given by $\mathbf{m}^{(\mathbf{r})}$, $\mathbf{n}^{(\mathbf{r})}$, and $\mathbf{I}^{(\mathbf{s})}$. The overlap is the projection of pattern $\mathbf{x}$ onto $\mathbf{y}$, so that two patterns are orthogonal if and only if their overlap is zero.

An important property of rank-R matrices, such as the connectivity matrix J, is that their nonzero eigenvalues coincide with the eigenvalues of the overlap matrix J^{ov} (Nakatsukasa, 2019) that is defined by the overlaps between pairs of connectivity patterns,

$$J_{rs}^{ov} = O\left(\mathbf{m}^{(\mathbf{s})}, \mathbf{n}^{(\mathbf{r})}\right), \tag{2.8}$$

for $r, s = 1, \ldots, R$. The eigenvalues of the connectivity matrix, and therefore of the overlap matrix, are an essential property to understand the dynamics of low-rank networks, as we show in section 4. It is often more convenient to calculate the eigenspectrum of the overlap matrix J^{ov}, of size $R \times R$, than of the connectivity matrix J, of size $N \times N$.

In a network with P populations, any pattern $\mathbf{x}$ of length N can be represented as a set of P subpatterns $\mathbf{x}^{(\mathbf{p})}$, for $p = 1, \ldots, P$, where each subpattern has length $\alpha_p N$ and includes the components of neurons belonging to population p. Figure 1 shows an example of a rank-two network with two populations, where the connectivity patterns can be split into two different subpatterns of equal size (green and purple). The overlap between two patterns can then be expressed as a weighted average of the overlaps between subpatterns:

$$O(\mathbf{x}, \mathbf{y}) = \sum_{p=1}^{P} \alpha_p O\left(\mathbf{x}^{(\mathbf{p})}, \mathbf{y}^{(\mathbf{p})}\right). \tag{2.9}$$

Even if the subpatterns are not orthogonal to each other (i.e. the overlap between two subpatterns is not zero), the patterns can be orthogonal to each

other when the subpattern overlaps cancel out. In the limit of large networks, the overlap between two subpatterns $\mathbf{x^{(p)}}$ and $\mathbf{y^{(p)}}$ is given by the expected value over the distribution of the loadings in the population:

$$O\left(\mathbf{x^{(p)}}, \mathbf{y^{(p)}}\right) = E\left[x^{(p)}y^{(p)}\right] = a_x^{(p)}a_y^{(p)} + \Sigma_{xy}^{(p)}. \tag{2.10}$$

In order to define the overlap matrix in terms of the statistics of the different gaussian populations, we define the matrix

$$\sigma_{n_r m_s}^{(p)} = \Sigma_{m_s n_r}^{(p)}. \tag{2.11}$$

The matrix $\boldsymbol{\sigma}_{mn}^{(p)}$ is an $R \times R$ whose entries contain the covariance between the connectivity patterns $\mathbf{m^{(r)}}$ and the $\mathbf{n^{(r)}}$ in population p. We call this matrix $\boldsymbol{\sigma}_{mn}^{(p)}$ a (reduced) covariance matrix, in an abuse of notation, because it is a subset of the covariance matrix $\Sigma^{(p)}$, and therefore it is not symmetric or positive definite. For example, for a rank-one network, $\boldsymbol{\sigma}_{mn}^{(p)}$ is just a scalar that can take any real value. For a rank-two network, $\boldsymbol{\sigma}_{mn}^{(p)}$ is a 2×2 matrix, whose entries are given by the four covariances $\sigma_{m_1 n_1}^{(p)}$, $\sigma_{m_1 n_2}^{(p)}$, $\sigma_{m_2 n_1}^{(p)}$, and $\sigma_{m_2 n_2}^{(p)}$.

Using equations 2.9 and 2.10, we can characterize the overlap matrix J^{ov} as a function of the statistics of the connectivity subpatterns,

$$J^{ov} = \sum_{p=1}^{P} \alpha_p \left(a_n^{(p)} a_m^{(p)^T} + \boldsymbol{\sigma}_{mn}^{(p)}\right), \tag{2.12}$$

where $a_n^{(p)}$ and $a_m^{(p)}$ are R-dimensional vectors whose entries correspond to the corresponding subset of elements in $a^{(p)}$ (see Figure 1C).

Similar to the covariance matrix $\boldsymbol{\sigma}_{mn}$ that measures the correlations between connectivity patterns $\mathbf{m^{(r)}}$ and $\mathbf{n^{(r)}}$, we define the covariance $\sigma_{\mathbf{nI}}$ between the connectivity patterns $\mathbf{n^{(r)}}$ and the constant external input $\mathbf{I}$ as a vector of length R, where each component is defined as

$$\sigma_{n_r I}^{(p)} = \Sigma_{n_r I}^{(p)} \tag{2.13}$$

for $r = 1, \ldots, R$. We assume that the input loadings and loadings of the left connectivity patterns are uncorrelated within each pattern, $\sigma_{m_r I}^{(p)} = 0$.

3 Dynamics in Gaussian Mixture Low-Rank Networks

In this section, we present three key properties of dynamics in mixture of gaussian low-rank networks: (1) in a network of rank R, dynamics can be characterized by R collective variables that form a dynamical system; (2) for

loadings drawn from gaussian mixture distributions, the dynamics can be further described as an effective circuit in which collective variables interact through gain-modulated effective couplings; and (3) with a sufficient number of populations, the resulting low-dimensional dynamics can approximate an arbitrary R-dimensional dynamical system.

Details of the derivations are provided in appendixes A and B.

3.1 Low-Dimensional Dynamics. In recurrent networks with low-rank connectivity, the dynamics of the trajectories $\mathbf{x}(t)$ are embedded in a linear subspace of dimension $R + N_{in}$ spanned by the left singular vectors $\mathbf{m^{(r)}}$ and the external input patterns $\mathbf{I^{(s)}}$ and can therefore be expressed as

$$x_i(t) = \sum_{r=1}^{R} \kappa_r m_i^{(r)} + \sum_{s=1}^{N_{in}} \kappa_{I_s} I_i^{(s)}. \tag{3.1}$$

Here κ_r and κ_{I_s} are collective variables that are obtained by projecting the activity $\mathbf{x}(t)$ on the patterns $\mathbf{m^{(r)}}$ and $\mathbf{I^{(s)}}$, that we assume are orthogonal to each other. Introducing the trajectory $\mathbf{x}(t)$ expressed in this new basis into equation 2.1, the dynamics of the collective variables are then given by the following dynamical system:

$$\tau \frac{d\kappa_r}{dt} = -\kappa_r + \kappa_r^{rec} \tag{3.2}$$

$$\tau \frac{d\kappa_{I_s}}{dt} = -\kappa_{I_s} + u_s(t)$$

$$\kappa_r^{rec} = \frac{1}{N} \sum_{i=1}^{N} n_i^{(r)} \phi \left(\sum_{s=1}^{N_{in}} I_i^{(s)} \kappa_{I_s} + \sum_{l=1}^{R} m_i^{(l)} \kappa_l \right). \tag{3.3}$$

We focus in the following on networks receiving a constant input, so that there is only one collective variable κ_I along the input dimension, the value of which is constant. The recurrent connectivity contributes to the dynamics of κ_r through the term κ_r^{rec}.

The dynamics of collective variables in equation 3.2 are valid for any finite-size, low-rank network, without any assumption on the values of pattern loadings. We next turn to networks where the pattern loadings are generated from specific distributions.

3.2 Dynamics in Multipopulation Networks. For low-rank networks in which pattern loadings are generated for each neuron from a gaussian mixture distribution, in the limit of large N, the dynamics in equation 3.2 can be expressed in terms of the statistics of pattern loadings over the populations and become (see appendix A)

$$\tau \frac{d\kappa_r}{dt} = -\kappa_r + \kappa_r^{rec} \tag{3.4}$$

$$\kappa_r^{rec} = \sum_{p=1}^{P} \alpha_p \left[a_{n_r}^{(p)} \left\langle \phi \left(\mu^{(p)}, \Delta^{(p)} \right) \right\rangle \right.$$

$$\left. + \left(\sigma_{n_r I}^{(p)} \kappa_I + \sum_{s=1}^{R} \sigma_{n_r m_s}^{(p)} \kappa_s \right) \left\langle \phi' \left(\mu^{(p)}, \Delta^{(p)} \right) \right\rangle \right]. \tag{3.5}$$

Here $\mu^{(p)}$ and $\Delta^{(p)}$ are the mean and variance of input to population p, given by

$$\mu^{(p)} = a_I^{(p)} \kappa_I + \sum_{s=1}^{R} a_{m_s}^{(p)} \kappa_s, \tag{3.6}$$

$$\Delta^{(p)} = \sigma_{I^2}^{(p)} \kappa_I^2 + \sum_{r=1}^{R} \sigma_{m_r^2}^{(p)} \kappa_r^2. \tag{3.7}$$

In equation 3.5, we used the gaussian integral notation:

$$\left\langle f(\mu, \Delta) \right\rangle = \int dx \, (2\pi)^{-\frac{1}{2}} e^{-x^2/2} f\left(\mu + \sqrt{\Delta} x \right). \tag{3.8}$$

The gaussian integral notation $\left\langle f(\mu, \Delta) \right\rangle$ represents the expected value of the random variable obtained after applying the function f to a random gaussian variable characterized by mean μ and variance Δ.

The factor $\left\langle \phi' \left(\mu^{(p)}, \Delta^{(p)} \right) \right\rangle$ in equation 3.5 corresponds to the average gain of neurons in population p in a given state, specified by the mean $\mu^{(p)}$ and variance $\Delta^{(p)}$ of the inputs to the population p. For each population, this average gain multiplies the covariances $\sigma_{m_l n_r}^{(p)}$ and $\sigma_{n_r I}^{(p)}$, and the corresponding average over populations defines an effective connectivity:

$$\tilde{\sigma}_{xy} = \sum_{p=1}^{P} \alpha_p \sigma_{xy}^{(p)} \left\langle \phi' \left(\mu^{(p)}, \Delta^{(p)} \right) \right\rangle. \tag{3.9}$$

The contributions of the first-order statistics $a_{n_r}^{(p)}$ to the recurrent dynamics are modulated by the average firing rate in population p and define an effective input:

$$\tilde{a}_{n_r} = \sum_{p=1}^{P} \alpha_p a_n^{(p)} \left\langle \phi \left(\mu^{(p)}, \Delta^{(p)} \right) \right\rangle. \tag{3.10}$$

When the effective connectivity and inputs are introduced into equation 3.4 the dynamics of a low-rank network with uncorrelated constant input take the simple form of an effective circuit of interacting collective variables:

$$\tau \frac{d\kappa_r}{dt} = -\kappa_r + \tilde{a}_{n_r} + \sum_{l=1}^{R} \tilde{\sigma}_{n_r m_l} \kappa_l. \tag{3.11}$$

Note that equation 3.11 describes the full nonlinear dynamics in the limit $N \to \infty$. Although the collective variables interact linearly through the effective connectivity and inputs, those depend implicitly on κ_r. The overall dynamics are therefore nonlinear, the nonlinearity being fully encapsulated in the effective inputs and couplings.

3.3 Universal Approximation of Low-Dimensional Dynamical Systems. By mapping the dynamics in equations 3.4 and 3.11 to a feedforward network with a single hidden layer and exploiting the universal approximation theorem (Cybenko, 1989; Leshno, Lin, Pinkus, & Schocken, 1993), we can show that a gaussian mixture network of rank R receiving a constant input is a universal approximator of R-dimensional dynamical systems (appendix B). More precisely, for a sufficient number of populations, the low-rank dynamics in equations 3.5 and 3.11 can approximate with arbitrary precision any R-dimensional dynamical system,

$$\frac{d\boldsymbol{\kappa}}{dt} = G(\boldsymbol{\kappa}), \tag{3.12}$$

defined by a vector field,

$$G(\{\kappa_r\}_{r=1\ldots R}) := (G_1(\{\kappa_r\}_{r=1\ldots R}), \ldots, G_R(\{\kappa_r\}_{r=1\ldots R})), \tag{3.13}$$

over an arbitrary finite domain $\{\kappa_r\}_{r=1\ldots R} \in [\kappa_r^{min}, \kappa_r^{max}]$. More specifically, this result requires that the vector field G is bounded and piecewise continuous and the transfer function is not a polynomial (see appendix B).

As an alternative to approximating any vector field over a bounded domain, we show that if the transfer function is bounded and monotonic, a rank-R network with multiple populations can approximate any vector field $G(\{\kappa_r\}_{r=1\ldots R})$ over the full domain of the collective variables, $\{\kappa_r\}_{r=1\ldots R} \in [-\infty, +\infty]$, with the restriction that the vector field follows asymptotic leaky dynamics for large input values:

$$\lim_{\kappa_s \to \pm\infty} \frac{\partial G_r}{\partial \kappa_{r'}}(\kappa_1, \ldots, \kappa_r) = -\delta_{rr'} \tag{3.14}$$

for any values $s, r, r' = 1, \ldots, R$, where G_r represents the rth component of the vector field as in equation 3.13 and δ_{ij} is the Kronecker delta. This stems from the fact that for large values of κ_r, the recurrent dynamics (see equation 3.5) saturate to a constant value.

Note that the universal approximation theorem does not state how many populations P are required to implement a given dynamical system and does not provide an algorithm for finding the statistics of the different populations.

4 Dynamics in Networks with a Single Population

Having shown that a rank R network with an arbitrary number of populations can approximate any R-dimensional dynamical system, we now illustrate how having a small number of populations in contrast limits the possible dynamics.

We focus first on the case of networks consisting of a single gaussian population. For simplicity, we focus on autonomous networks with zero-mean connectivity patterns. Specifically, we show that independent of their rank, the range of dynamics such networks can implement is restricted. This case was previously studied for connectivities that combined a rank-one or rank-two structure and a full-rank random component (Mastrogiuseppe & Ostojic, 2018; Schuessler, Dubreuil et al., 2020). Here we provide an overview of those results and extend them to single-population networks of arbitrary rank. The fact that we focus on networks whose connectivity is low-rank allows us to provide a deeper analysis of the dynamics.

In vectorial form, assuming zero-mean connectivity patterns, the collective dynamics in equation 3.4 for one population read

$$\tau \frac{d\boldsymbol{\kappa}}{dt} = -\boldsymbol{\kappa} + \langle \phi'\left(0, \boldsymbol{\kappa}^T \boldsymbol{\kappa}\right)\rangle \boldsymbol{\sigma}_{mn}\boldsymbol{\kappa}, \tag{4.1}$$

where we used the vector of collective variables $\boldsymbol{\kappa} \in \mathcal{R}^R$ and the $R \times R$ covariance matrix $\boldsymbol{\sigma}_{mn}$ as defined in equation 2.11, which is equal to the overlap matrix, equation 2.12, in the case of zero-mean connectivity patterns. In the following analysis, we show that the eigenvalues of the covariance matrix $\boldsymbol{\sigma}_{mn}$, which for $N \to \infty$ are identical to the eigenvalues of the connectivity matrix, determine the dynamics in collective space. Schuessler, Dubreuil et al. (2020) performed a similar analysis for networks with random connectivity and rank-one and rank-two perturbations.

4.1 Fixed Points. The fixed points of equation 4.1 are given by

$$\boldsymbol{\kappa}_0 = \langle \phi'\left(0, \boldsymbol{\kappa}_0^T \boldsymbol{\kappa}_0\right)\rangle \boldsymbol{\sigma}_{mn}\boldsymbol{\kappa}_0. \tag{4.2}$$

For $\phi(x) = \tanh(x)$, the trivial point $\kappa_0 = 0$ is always a solution. There might, however, be nontrivial fixed points depending on the eigenvalues of the covariance matrix σ_{mn}. The covariance matrix can have up to R eigenvalues, which we denote λ_r, with associated eigenvectors u_r. Each real and nondegenerate eigenvalue λ_r of the covariance σ_{mn} generates a fixed point $\kappa_0^{(r)} = \rho_r u_r$, where ρ_r is the radial distance of the fixed point along the direction set by the eigenvector u_r. Introducing this parameterization of the fixed points in equation 4.2, we obtain the following implicit equation for the value ρ_r:

$$1 = \lambda_r \langle \phi'(0, \rho_r^2) \rangle. \tag{4.3}$$

The gain factor $\langle \phi'(0, \rho_r^2) \rangle$ is bounded between 0 and 1 for the transfer function $\phi(x) = \tanh x$. Therefore, eigenvalues $\lambda_r > 1$ generate two nontrivial fixed points, symmetrically located around the origin (see Figures 2A to 2D, bottom row, for a rank-one example). Smaller eigenvalues do not generate any nontrivial fixed point (Figures 2A to 2D, first row).

In order to determine the stability of the fixed points, we linearize the dynamics and obtain the Jacobian S_r at the fixed point corresponding to the eigenvalue λ_r of σ_{mn} (see appendix C),

$$S_r = -I + \frac{1}{\lambda_r}\sigma_{mn} + \langle \phi'''(0, \rho_r^2) \rangle \lambda_r \rho_r^2 u_r u_r^T, \tag{4.4}$$

where I denotes the $R \times R$ identity matrix. The eigenvalues of S_r determine the stability of the fixed points: if any positive eigenvalue exists, the dynamics will diverge away from the fixed point in the direction of the corresponding eigenvector of S_r. Negative eigenvalues correspond to attractive modes of the dynamics around the fixed point. If all eigenvalues of the stability matrix are negative, the fixed point is stable.

When the eigenvectors of the matrix σ_{mn} are orthogonal to each other (see Figures 3A to 3D), the R eigenvalues of the matrix S_r, denoted as $\gamma_{r'}$ for $r' = 1, \ldots, R$, can be calculated analytically: the eigenvalue $\gamma_{r'}$ has an associated eigenvector equal to the eigenvector $u_{r'}$ of the covariance matrix σ_{mn} and reads

$$\gamma_{r'} = -1 + \frac{\lambda_{r'}}{\lambda_r} + \langle \phi'''(0, \rho_r^2) \rangle \lambda_r \rho_r^2 \delta_{rr'}. \tag{4.5}$$

Remarkably, the eigenvalues of the Jacobian around any nontrivial fixed point are therefore directly determined by the eigenvalues of connectivity and covariance matrices (Schuessler, Dubreuil et al., 2020). If $r' = r$, the two first terms cancel out, and the third term is always negative (see appendix C). This implies that all nontrivial fixed points are stable in the direction u_r
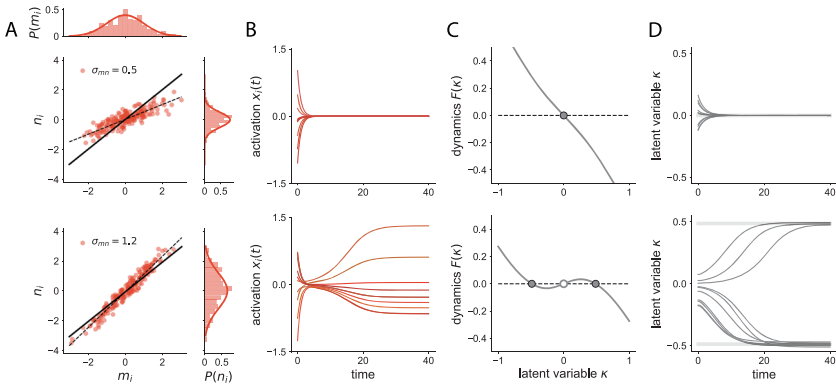
Figure 2: Dynamics in rank-one networks with a single gaussian population. (A) Scatter plot of the loadings of left singular vectors $m_i^{(r)}$ and right singular vectors $n_i^{(r)}$. Top: Covariance σ_{mn}, indicated by the slope of the dashed line, below the critical value for nontrivial fixed points (solid line). Bottom: Covariance σ_{mn} beyond the critical value. (B) Dynamics of the activation variable $x_i(t)$ of 10 units in the network for the two different networks initialized at random values. The network with σ_{mn} larger than one (bottom) converges to a heterogeneous fixed point, while the other one decays to zero. (C) One-dimensional dynamics corresponding to the right-hand side of equation 4.1. Filled dots correspond to stable fixed points. For a weak covariance between connectivity patterns (top), the trivial fixed point is the only fixed point. For a strong covariance (bottom), the recurrent connectivity generates two nontrivial stable fixed points. (D) Evolution of the collective variable κ as a function of time in a finite-size network, defined as the projection of the activity $\mathbf{x}(t)$ onto the connectivity pattern $\mathbf{m}$. Each curve corresponds to a different realization of the random connectivity matrix. $N = 1000$, top row: $\sigma_{n^2} = 0.34$; bottom row $\sigma_{n^2} = 1.52$.

that points toward the origin. However, if there are other nontrivial fixed points corresponding to eigenvalues $\lambda_{r'} > \lambda_r$ of $\boldsymbol{\sigma}_{mn}$, the fixed point $\boldsymbol{\kappa}_0^{(r)}$ is destabilized in the directions of the eigenvectors with larger eigenvalues.

When the eigenvectors of $\boldsymbol{\sigma}_{mn}$ are not orthogonal to each other, the eigenvectors of $\boldsymbol{\sigma}_{mn}$ are not eigenvectors of the linear stability matrix S_r. However, the eigenvalues of S_r are still given by equation 4.5 (see appendix D), so that the same stability properties hold (see appendix D): every fixed point is stable in the direction toward the origin, and the fixed point in the direction given by the largest eigenvalue is stable while the other ones become unstable.

In summary, if all eigenvalues of the covariance matrix are real and nondegenerate, only the pair of nontrivial fixed points corresponding to the largest eigenvalue is stable. All the other nontrivial fixed points of the dynamics are saddle points. This implies that low-rank networks
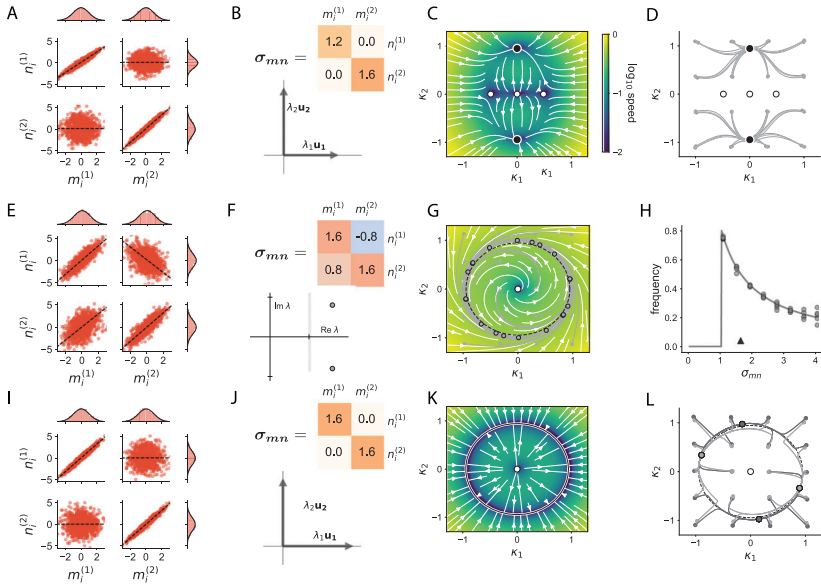
Figure 3: Dynamics in rank-two networks with a single gaussian population. (A) Scatter plot of the loadings of left singular vectors $m_i^{(r)}$ and right singular vectors $n_i^{(r)}$. (B) Covariance matrix σ_{mn} of the population (top), and its eigenvectors (bottom). (C) Vector field corresponding to the mean-field dynamics in the plane $\kappa_1 - \kappa_2$ of collective variables (see equation 4.1). The color map represents the speed of the dynamics, defined as the norm of vector $\frac{d\kappa}{dt}$, in different points of the collective space. Two nontrivial fixed points are generated in the direction of each eigenvector. Black dots correspond to stable fixed points, while white dots are unstable or saddle points. The pair of fixed points corresponding to the largest eigenvalue is stable. (D) Finite-size simulations of the dynamics. Three different connectivity realizations are shown from each initial condition. $N = 1000$. (E–G) Similar to panels A to C for a network with complex eigenvalues (overlap matrix given by equation 4.6). The network generates a limit cycle. Gray curves in panel G correspond to trajectories from finite-size networks. The dots represent the final state after a fixed time elapsed. (H) Frequency of the limit cycle for different values of the symmetric part of the connectivity σ and fixed imaginary part $\sigma_\omega = 0.8$. The dots show the numerically estimated frequency of oscillations in finite-size simulations for five different network realizations. The line corresponds to equation 4.10. The triangle indicates the parameter σ used in panels E to G. (I–L) Similar to panels A to D for a network with degenerate eigenvalues: any vector in the plane spanned by vectors $\mathbf{m}^{(1)}$ and $\mathbf{m}^{(2)}$ is an eigenvector of the connectivity. This symmetry leads to a continuous attractor in the mean-field dynamics. In finite size simulations (two connectivity matrix realizations shown in panel L, in different shades of gray, with filled points corresponding to the stable fixed points) the continuous attractor corresponds to a slow manifold on which usually two stable fixed points lie. Parameters: $\sigma_{n_1^2} = 1.24$, $\sigma_{n_2^2} = 1.63$.

consisting of a single gaussian population can have at most two stable fixed points independent of their rank.

4.2 Limit Cycles. Complex eigenvalues of the covariance matrix σ_{mn}, if they exist, always appear in conjugate pairs. They lead to spiral dynamics around the origin, in the plane spanned by the real and imaginary part of the corresponding eigenvectors. If the real part of the complex eigenvalues is smaller than unity, $\text{Re}(\lambda_r) < 1$, the spiral dynamics decay back to the origin. Otherwise, if $\text{Re}(\lambda_r) > 1$, there is a limit cycle on the plane, around the origin. Similar to the case with only real eigenvalues of the covariance matrix, if the real part of the complex eigenvalue is larger than the real part of any other eigenvalue of σ_{mn}, any trajectory will converge to the plane defined by the real and imaginary parts of the corresponding eigenvectors. On this plane, we then find that the limit cycle is stable.

To illustrate this case, we consider a rank two network with a covariance matrix of the form

$$\sigma_{mn} = \begin{pmatrix} \sigma & -\sigma_\omega \\ \sigma_\omega & \sigma \end{pmatrix}, \tag{4.6}$$

which has eigenvalues $\sigma \pm i\sigma_\omega$. Figures 3E and 3F show an example of a network with such connectivity.

We can then write the equations for a rank-two network in polar form. Any state $\kappa = [\kappa_1, \kappa_2]$ in a rank-two network can be mapped to the radial distance ρ and an orientation θ using the mapping $\kappa_1 := \rho \cos\theta$ and $\kappa_2 := \rho \sin\theta$. The dynamics in equation 4.1 become

$$\tau \frac{d\rho}{dt} = -\rho + \rho\sigma \left\langle \phi'\left(0, \rho^2\right) \right\rangle, \tag{4.7}$$

$$\tau \frac{d\theta}{dt} = \sigma_\omega \left\langle \phi'\left(0, \rho^2\right) \right\rangle. \tag{4.8}$$

When the real part σ of the eigenvalues is larger than one, the flow in the radial direction cancels at a value ρ_0 given by equation 4.3, which yields

$$\sigma^{-1} = \left\langle \phi'\left(0, \rho_0^2\right) \right\rangle. \tag{4.9}$$

Based on equation 4.7, we observe that any perturbation in the plane away from the limit cycle makes the radial component ρ go back to ρ_0. The limit cycle is therefore stable, as shown in Figure 3G.

Introducing this result into equation 4.8, we obtain that the oscillations of the limit cycle are generated at a frequency

$$\omega_{LC} = \frac{\sigma_\omega}{\sigma}. \tag{4.10}$$

In this analysis, equation 4.10 is derived for the particular covariance matrix $\boldsymbol{\sigma}_{mn}$ in equation 4.6, which is the sum of an isotropic matrix (proportional to the identity) and an antisymmetric matrix. However, this equation is valid more generally for any connectivity matrix with a pair of complex eigenvalues (see appendix D). When the covariance matrix is not antisymmetric but still has complex eigenvalues, the limit cycle is no longer a circle but resembles an ellipse, while the frequency of oscillation appears to still be given by equation 4.10.

Complex eigenvalues can be combined with real eigenvalues in networks with rank larger than two. The same stability properties are kept: only the fixed point or the limit cycle generated by the eigenvector with the largest real part is stable. The unstable fixed points or limit cycles remain attractive within the dimensions spanned by the corresponding eigenvector (see Figure 9 in appendix D) for an example of a rank-three network combining an unstable limit cycle and two stable fixed points).

4.3 Slow Manifolds. When the covariance matrix $\boldsymbol{\sigma}_{mn}$ has degenerate eigenvalues, low-rank RNNs can lead to phenomena other than discrete fixed points or limit cycles. As an example of degenerate eigenvalues, we study the network dynamics when the covariance matrix $\boldsymbol{\sigma}_{mn}$ is diagonal:

$$\boldsymbol{\sigma}_{mn} = \sigma_{mn}\boldsymbol{I}. \tag{4.11}$$

This covariance matrix has one single real eigenvalue σ_{mn}, which is degenerate, since it has R linearly independent eigenvectors. Introducing the covariance matrix in equation 4.11 into the dynamics in equation 4.1, we obtain the fixed-point equation

$$\boldsymbol{\kappa}_0 = \left\langle \phi'\left(0, \boldsymbol{\kappa}_0^T\boldsymbol{\kappa}_0\right)\right\rangle \sigma_{mn}\boldsymbol{\kappa}_0. \tag{4.12}$$

To solve the fixed-point equation, as in the previous section, we use the ansatz $\boldsymbol{\kappa}_0 = \rho_0\boldsymbol{u}_{\kappa_0}$, where $\boldsymbol{u}_{\kappa_0}$ is an arbitrary unitary vector in collective space. Introducing the ansatz in the fixed-point equation 4.12, we find that there is a nontrivial solution given implicitly by the scalar equation $\left\langle \phi'\left(0, \rho_0^2\right)\right\rangle = \sigma_{mn}^{-1}$, which is independent of the particular direction $\boldsymbol{u}_{\kappa_0}$. Furthermore, we find that the fixed point is stable in the direction $\boldsymbol{u}_{\kappa_0}$. Therefore, in the mean-field limit given by equation 4.1, this degenerate connectivity leads to a continuous manifold of attractive states that are at an equal distance ρ_0 away from the origin. In the case of rank-two connectivity, this degenerate covariance matrix leads to a stable ring attractor (see Figures 3I to 3K), and in rank-R, to a stable R-spherical attractor.

In finite-size simulations, the sampling of random loadings introduces spurious correlations in the matrix $\boldsymbol{\sigma}_{mn}$, breaking the degeneracy of the eigenvalues. As a consequence, only a small number of points on the

continuous attractor predicted by the mean-field theory give rise to actual fixed points. While the rest of the points on the predicted continuous attractor are not fixed points of the finite-size network, the dynamics around them are typically slow. More specifically, any trajectory of activity quickly converges toward the predicted continuous attractor and then slowly evolves along it until it reaches a fixed point (see Figure 3L) (Mastrogiuseppe & Ostojic, 2018). In finite-size networks, the continuous attractor predicted by the mean-field analysis therefore gives rise to a low-dimensional manifold in state space, along which the dynamics are slow.

When degenerate and nondegenerate real and complex eigenvalues are combined, the global stability appears to be given by the criterion in equation 4.5: each eigenvalue generates its corresponding nontrivial dynamics (fixed points, continuous attractors or limit cycle) independently. The stability of these dynamical phenomena depends on the global eigenspectrum: the eigenvalues with the largest real part generate stable attractors, while the other eigenvalues lead to repellers.

4.4 Summary. In a low-rank network consisting of a single gaussian population, the possible nontrivial steady states are a pair of fixed points, a limit cycle, or a continuous attractor that gives rise to a small number of fixed points in finite networks. On top of this limited range of stable solutions, increasing the rank leads to additional unstable fixed points and limit cycles that can potentially be used to control the dynamics, a point we do not further explore here. We instead proceed to show that increasing the number of gaussian populations allows networks to implement a larger range of stable dynamics.

5 Dynamics in Networks with Multiple Populations

As described in the previous section, a major limitation of rank-*R* networks consisting of a single gaussian population is that they cannot give rise to more than two stable fixed points, symmetrically arranged around the origin. We next show that networks consisting of several gaussian populations can exhibit a larger number of stable fixed points. We specifically describe two mechanisms by which multiple fixed points can be generated and controlled and show that classical Hopfield networks (Hopfield, 1982) correspond to a particular limit of gaussian-mixture low-rank networks.

5.1 Nonlinear Gain Control. The first mechanism for generating multiple fixed points is based on having several populations that reach saturation in different regions of the collective space. The local dynamics in these different regions are then controlled solely by the statistics of the nonsaturated populations.

For concreteness, we consider a rank-one network consisting of two populations, defined by different statistics of pattern loadings. Within
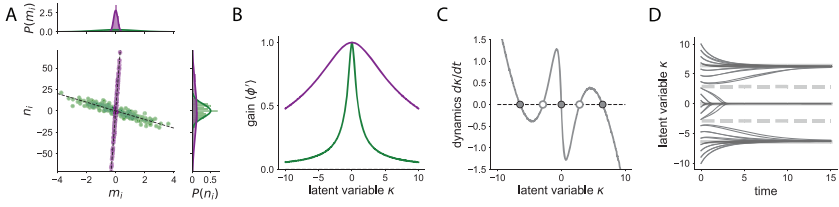
Figure 4: Nonlinear gain control in rank-one networks with multiple populations. (A) Scatter plot between the components of the connectivity patterns m_i and n_i in a rank-one network with two gaussian populations, shown in green (negatively correlated population) and purple (positively correlated population). (B) Average gain within the two populations (green and purple lines) at different states of the collective variable κ. The green population (large variance $\sigma_{m^2}^{(1)}$) saturates at values much closer to the origin than the purple population (low variance $\sigma_{m^2}^{(2)}$). Therefore, for large values of κ, the purple population has a stronger influence on the dynamics. (C) Mean-field dynamics generated by the two-population statistics. Three stable fixed points (filled gray dots) emerge in the 1D recurrent dynamics. Close to the origin κ_0, the fixed point is stable because the green population dominates, and it has a negative correlation $\sigma_{mn}^{(1)} < 0$. Therefore, the origin is stable. At values of κ far from the origin, the purple population dominates, and it creates nontrivial stable fixed points. (D) Dynamics of the collective variable κ in a network with $N = 1000$ units, initiated at different initial values. The dynamics converge to one of the three stable fixed points. Parameters: $\sigma_{mn}^{(1)} = -10$, $\sigma_{mn}^{(2)} = 4.5$, $\sigma_{m^2}^{(1)} = 1.98$, $\sigma_{m^2}^{(2)} = 0.02$, $\sigma_{n^2}^{(1)} = 59.5$, $\sigma_{n^2}^{(2)} = 1020$, and $\alpha_1 = \alpha_2 = 0.5$.

population p, for $p = 1, 2$, the joint distribution of n and m values over neurons is specified by a 2×2 covariance matrix $\Sigma^{(p)}$. For simplicity, we take the mean of the distribution to be zero. In the two-dimensional loading space defined by m and n, the two populations correspond to different gaussian clusters, both centered at zero but with different shapes and orientations (see the green and purple dots in Figure 4A).

The dynamics of the collective variable κ in equation 3.11 read:

$$\tau \frac{d\kappa}{dt} = -\kappa + \tilde{\sigma}_{mn}\kappa, \tag{5.1}$$

with the effective feedback $\tilde{\sigma}_{mn}$ defined as

$$\tilde{\sigma}_{mn} = \frac{1}{2}\sigma_{mn}^{(1)}\left\langle \phi'\left(0, \kappa^2\sigma_{m^2}^{(1)}\right)\right\rangle + \frac{1}{2}\sigma_{mn}^{(2)}\left\langle \phi'\left(0, \kappa^2\sigma_{m^2}^{(2)}\right)\right\rangle. \tag{5.2}$$

This effective feedback $\tilde{\sigma}_{mn}$ is set by the average of covariances $\sigma_{mn}^{(p)}$ for each population p, weighted by the gain of the population. Low gain implies

that the population is at a saturated state. The parameter $\sigma_{m^2}^{(p)}$ controls the range at which a population saturates as the collective variable κ increases. If the two populations have different variances $\sigma_{m^2}^{(p)}$, their gains will vary differently with κ (see Figure 4B). If, moreover, the different populations have covariances $\sigma_{mn}^{(p)}$ of different signs, the total effective feedback will vary strongly at different ranges of κ, while this is not the case in networks with uniform populations or a single one. Therefore, by manipulating the variance of the **m** connectivity pattern within each population and the overlap between the left and right connectivity patterns, it is possible to generate more flexible dynamics.

In particular, the network can have three stable fixed points: one at the origin and a pair of symmetrical nontrivial fixed points. First, the origin $\kappa = 0$ is always a fixed point of dynamics in equation 5.1. The origin is, moreover, a stable fixed point if the effective feedback at zero, which is given by $\frac{1}{2}\left(\sigma_{mn}^{(1)} + \sigma_{mn}^{(2)}\right)$, is smaller than 1. Second, at large values of κ, the effective feedback $\tilde{\sigma}_{mn}$ should be positive to cancel the contribution of the leaky term $-\kappa$ and generate a nontrivial fixed point. Therefore, one of the populations, which we define to be the first one ($p = 1$), must have a strong, negative overlap, $\sigma_{mn}^{(1)} < 2 - \sigma_{mn}^{(2)} < 0$. Given equation 5.2, this implies that the gain of the positively correlated population two should be large, whereas the gain of the negatively correlated population one should be close to zero. A small gain is achieved in the first population by having a large value $\sigma_{m^2}^{(1)}$, so that the second condition reads $\sigma_{m^2}^{(1)} \gg \sigma_{m^2}^{(2)}$. Figures 4C and 4D show the dynamics of such a network given by the mean-field equation and in finite-size networks.

More generally, with more than two populations, this mechanism can be extended to produce a larger number of stable fixed points in rank-one networks. The two key components of this mechanism are (1) an independent control of the gain of the different populations, so that the contribution of each population to the effective feedback takes place in different ranges of the collective variable κ, and (2) covariances $\sigma_{mn}^{(p)}$ of different signs, so that the effective feedback can flexibly take both positive and negative values in different ranges of κ. These mechanisms can also be applied to networks with rank higher than one. In that case, the overlap between loadings is given by a matrix $\sigma_{mn}^{(p)}$ instead of a scalar, while the gain of each population is a scalar value. Populations with different covariance matrices and gains that vary at different ranges of the collective variables are able to generate multiple fixed points in different regions of the collective space, or combinations between stable limit cycles and stable fixed points (Dubreuil, Valente, Beiran, Mastrogiuseppe, & Ostojic, 2020).

5.2 Symmetries in Loading Space. A second mechanism for generating multiple fixed points is to exploit symmetries in the distribution of loadings

$P(\underline{m}, \underline{n})$. Such a symmetry in the connectivity induces a symmetry in the dynamics of the collective variables. In consequence, if a network generates one nontrivial stable fixed point, additional stable fixed points appear at symmetric points in the collective space.

We focus here on networks where the overlap between the connectivity patterns is given by the nonzero means of the loadings, which is complementary to the previous section where the connectivity patterns had zero mean and the recurrent dynamics is determined by the covariances between the loadings. We introduce a symmetry in the distribution of loadings by arranging the means of the loadings accordingly. Note, however, that symmetrical distributions of loadings can also be generated in the zero-mean case.

As an illustration, we consider first a rank-two network, with units evenly split into P populations. In each population, the loadings $m_1^{(p)}$, $m_2^{(p)}$, $n_1^{(p)}$, $n_2^{(p)}$ have a different set of means $a_{m_1}^{(p)}$, $a_{m_2}^{(p)}$, $a_{n_1}^{(p)}$, $a_{n_2}^{(p)}$, and the covariances $\sigma_{m_r n_s}^{(p)}$ are zero. The variance, of the loadings, σ_{m^2} and σ_{n^2}, are identical in all populations. As a consequence, different populations correspond to clusters of identical spherical shape, but centered at different points in the four-dimensional loading space.

We specifically arrange the means of the different populations (centers of the different clusters) symmetrically at the vertices of a regular polygon in the planes of loadings $m_1 - m_2$ and $n_1 - n_2$:

$$a_{m_1}^{(p)} = R_m \cos\left(\frac{2\pi p}{P}\right), \quad a_{m_2}^{(p)} = R_m \sin\left(\frac{2\pi p}{P}\right), \tag{5.3}$$

$$a_{n_1}^{(p)} = R_n \cos\left(\frac{2\pi p}{P}\right), \quad a_{n_2}^{(p)} = R_n \sin\left(\frac{2\pi p}{P}\right), \tag{5.4}$$

where p is the population index, $p = 1 \ldots P$ for $P > 2$. The radial distance R_m is fixed so that the patterns $\mathbf{m}^{(1)}$ and $\mathbf{m}^{(2)}$ have unit variance, while the free parameter R_n controls the overlap between the connectivity patterns. Figure 5A shows an example with six populations, $P = 6$. This distribution has a discrete rotational symmetry of order P, since rotations of angle $2\pi/P$ in the planes $m_1 - n_2$ and $m_2 - n_1$ leave the distribution unchanged.

When the mean-field description in equation 3.4 is used, the dynamics of the two collective variables now read

$$\tau \frac{d\kappa_1}{dt} = -\kappa_1 + \frac{1}{P} \sum_{p=1}^{P} a_{n_1}^{(p)} \left\langle \phi\left(a_{m_1}^{(p)}\kappa_1 + a_{m_2}^{(p)}\kappa_2, \sigma_m^2\left(\kappa_1^2 + \kappa_2^2\right)\right)\right\rangle, \tag{5.5}$$

$$\tau \frac{d\kappa_2}{dt} = -\kappa_2 + \frac{1}{P} \sum_{p=1}^{P} a_{n_2}^{(p)} \left\langle \phi\left(a_{m_1}^{(p)}\kappa_1 + a_{m_2}^{(p)}\kappa_2, \sigma_m^2\left(\kappa_1^2 + \kappa_2^2\right)\right)\right\rangle. \tag{5.6}$$
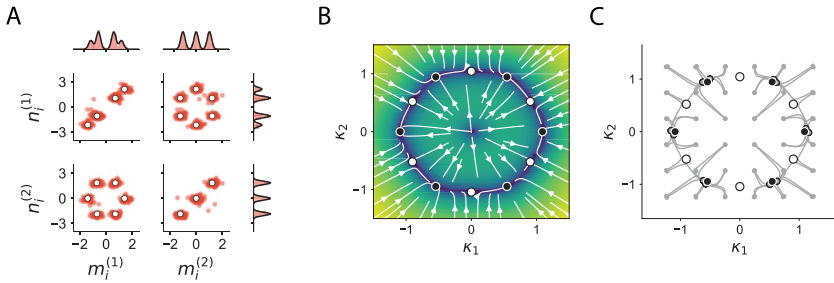
Figure 5: Discrete rotational symmetry in rank-two networks with multiple populations. (A) Scatter plot between the components of the connectivity patterns in a rank-two network. The network consists of six populations, with centers located on the vertices of a regular hexagon. The distribution is invariant to rotations of an angle $2\pi/6$ in the m_1–n_2 and m_2–n_1 planes. (B) Mean-field dynamics of the network, the color map represents the speed of the dynamics Q (blue: slow dynamics, yellow: fast dynamics). The hexagonal symmetry in the loadings produces a solution with hexagonal symmetry, with six stable fixed points (black dots) symmetrically arranged along a ring. Saddle points (white dots) appear between the stable fixed points. (C) Trajectories of the collective variables in finite-size simulations, initiated at different initial conditions. All trajectories converge to one of the six stable fixed points. Two different network realizations are shown for each initial condition. Parameters: centers arranged as in equations 5.3 and 5.4 where $p = 6$ and $R_n = 1.5$. Variance $\sigma_{n^2} = 0.2$, equal in each population. Network size $N = 1000$.

Given the symmetry in the distribution, if we identify one nontrivial stable fixed point, there will be at least $P - 1$ other fixed points with the same stability. When the direction given by $\kappa_2 = 0$ is focused on, the velocity in the κ_2 direction, given by equation 5.6 is always zero due to the symmetry in the distribution. Therefore, we obtain a fixed-point equation for κ_1 on the $\kappa_2 = 0$ direction using equation 5.5:

$$\kappa_1 = \frac{1}{P} \sum_{p=1}^{P} R_n \cos\left(\frac{2\pi p}{P}\right) \left\langle \phi\left(R_m \cos\left(\frac{2\pi p}{P}\right)\kappa_1, \sigma_m^2 \kappa_1^2\right)\right\rangle. \tag{5.7}$$

The right-hand side is a sum of P monotonically increasing bounded functions of κ_1. If the slope at the origin is larger than one, the right-hand side will intersect with the function κ_1 at a nontrivial point. The slope of the right-hand side at the origin, obtained by differentiating the right-hand side with respect to κ_1 and evaluating at $\kappa_1 = 0$, is $\frac{1}{2}R_n R_m$, so that a condition for a nontrivial fixed point is

$$R_n R_m > 2. \tag{5.8}$$

Because of the symmetry, if $R_m R_n > 2$, there are at least P stable fixed points arranged symmetrically on a circle (see Figures 5A to 5C for a network with $P = 6$ populations and Figures 6A to 6C for a network with $P = 4$ populations). If the number of population pairs is odd, there are $2P$ stable fixed points symmetrically arranged on a circle, because there is also a symmetry with respect to the origin, imposed by the symmetry in the transfer function. Otherwise, if P is even, P stable fixed points are generated by the network.

Symmetrical arrangements of multiple populations can also be used in higher R-rank networks to obtain multiple stable fixed points located on an R-dimensional sphere. For example, in rank-three networks, we consider eight populations whose centers are arranged at the vertices of a cube. The centers of the eight populations in the three-dimensional space of loadings $m^{(r)}$, for $r = 1, 2, 3$, correspond to the vertices of a cube with side $2R_m$, so that

$$\left(a_{m_1}^{(p)}, a_{m_2}^{(p)}, a_{m_3}^{(p)} \right) = \left(\pm R_m, \pm R_m, \pm R_m \right). \tag{5.9}$$

Populations $p = 1, \ldots, 8$ correspond to one of the eight different possible combinations of the sign. The variances of the loadings, σ_{m^2}, are identical in all populations. The value of R_m is fixed so that the norm of each connectivity pattern $\mathbf{m^{(r)}}$ is N.

The centers of the $n^{(r)}$ loadings follow the same configuration, at the vertices of a cube of side $2R_n$:

$$\left(a_{n_1}^{(p)}, a_{n_2}^{(p)}, a_{n_3}^{(p)} \right) = \left(\pm R_n, \pm R_n, \pm R_n \right), \tag{5.10}$$

where each population p corresponds to the same combination of signs as for the m loadings, so that

$$\text{sgn} \left(a_{m_r}^{(p)} \right) = \text{sgn} \left(a_{n_r}^{(p)} \right), \tag{5.11}$$

with the collective index $r = 1, 2, 3$ and the population index $p = 1, \ldots, 8$. The value R_n is, as in the previous case, a free parameter that controls the overlap between connectivity patterns. This configuration is shown in Figures 6D, 6E, 6G, and 6H for two different values of R_n. This distribution exhibits a cubic symmetry in the loading space $m_1 - m_2 - m_3$ and in space $n_1 - n_2 - n_3$. Thus, if we identify a nontrivial fixed point, these symmetries require the existence of symmetric solutions in the collective space. Inspecting the direction $\kappa_2 = \kappa_3 = 0$ in the dynamics, we obtain a criterion for having a nontrivial stable fixed point:

$$\kappa_1 = \frac{1}{8} \sum_{p=1}^{8} a_{n_1}^{(p)} \left\langle \phi \left(a_{m_1}^{(p)} \kappa_1, \sigma_m^2 \kappa_1^2 \right) \right\rangle. \tag{5.12}$$
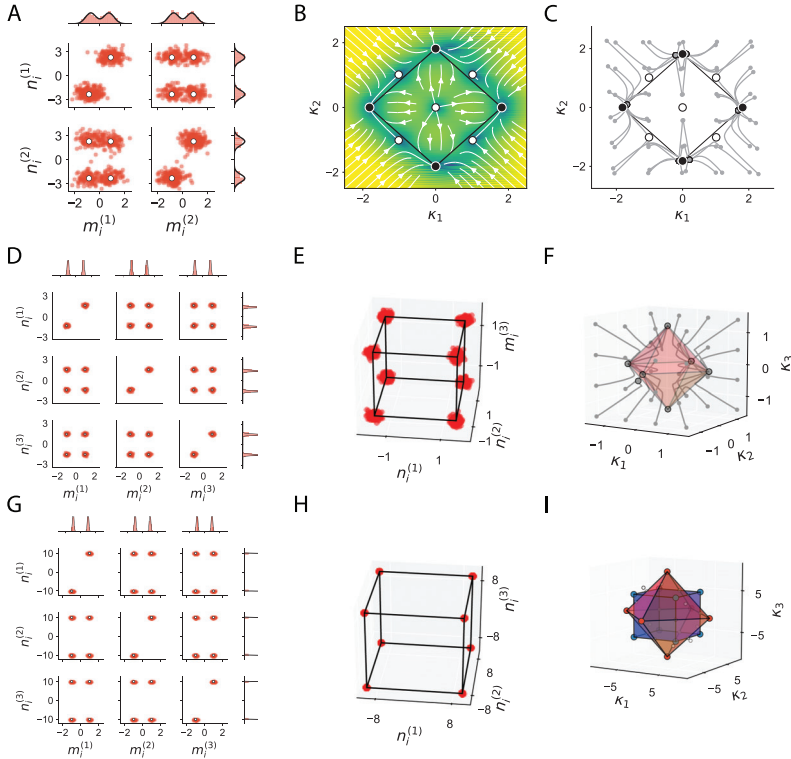
Figure 6: Dynamics in rank-R networks with discrete symmetry in multiple populations: Hopfield networks. (A) Scatter plot between the entries of left singular vectors m_i and right singular vectors n_i in a rank-two network with four populations following equations 5.3 and 5.4, with $P = 2$, $R_n = 2.3$, and $\sigma_{n_i^2} = 0.5$. (B) Corresponding mean-field dynamics. The color map represents the speed of the dynamics, defined as the norm of vector $\frac{d\kappa}{dt}$ (blue: slow dynamics, yellow: fast dynamics). Four stable fixed points emerge, arranged in a square. (C) Trajectories starting at different initial conditions in a finite-size network. Each initial condition shows trajectories for two network realizations. (D) Analogous to panel A in a rank-three network with loadings arranged as in equations 5.9 and 5.10. (E) The populations are arranged at the vertices of a cube in loading space. $R_n = 2.1$. (F) Dynamics of the collective variables. Six stable fixed points (gray dots) emerge, arranged at the vertices of a dodecahedron (dual polygon of the cube, highlighted in red for visual purposes). Gray lines correspond to the trajectories of finite-size networks, initialized at different points in state-space. (G–I) Same as in panels D to F but for a network whose populations have larger mean values, $R_n = 7$. For such large values, spurious fixed points that are proportional to the combinations of the three stored patterns ($\pm m_1 \pm m_2 \pm m_3$) also become stable. Therefore, apart from the six fixed points in a octahedron (red polygon), eight other spurious fixed points appear arranged in a cube (blue polygon). Network size $N = 1000$.

Equation 5.12 has a nontrivial solution, which is always stable if $R_n R_m > 1$. When this solution exists, applying a rotation of $\pi/2$ in the $m_1 - m_2$ plane and in the $m_1 - m_3$, it is possible to determine the other five stable fixed points that are generated by the symmetry (see Figure 6F). These stable fixed points are arranged in the collective space at the vertices of an octahedron, the dual polyhedron of the cube (the dual of a polyhedron A is the polyhedron B where the vertices of A correspond to the edges of B). When symmetry principles are applied, the middle point of each triangular face of the octahedron is also a fixed point. However, the stability of this fixed point depends on the overlap $R_n R_m$. If $R_n R_m$ is larger than one but low, these fixed points are saddle points (see Figure 6F). Beyond a critical value of $R_n R_m$, these fixed points also become stable. This second set of fixed points consists of eight points arranged on a cube (see Figure 6I, blue dots).

In general, any K-dimensional discrete symmetry in the loadings will generate a dynamical system with stable fixed points on a K-dimensional sphere, arranged with the symmetry of the dual polytope. A regular polytope is defined as the generalization of a regular polyhedron generalized to more than three dimensions.

5.3 Relation to Hopfield Networks. Classical Hopfield networks (Hopfield, 1982) storing $R \ll N$ patterns can be seen as a particular limit of gaussian-mixture low-rank networks, where multiple stable fixed points are generated based on symmetries in the connectivity. A Hopfield network is designed to store R binary patterns $m_i^{(r)} = \pm m$, where for every neuron, the sign of the entry in each pattern is generated randomly and m is a scalar parameter. A Hopfield network storing these R patterns is defined as a recurrent network with connectivity matrix

$$J_{ij}^{Hopfield} = \sum_{r=1}^{R} m_i^{(r)} m_j^{(r)}. \tag{5.13}$$

Such a configuration corresponds to a specific type of low-rank matrix and can be mapped onto gaussian-mixture low-rank networks. A first specific property of Hopfield networks (see equation 5.13) is that the connectivity is symmetric, so that the left and right connectivity patterns are proportional to each other,

$$\mathbf{m^{(r)}} = c\mathbf{n^{(r)}}, \tag{5.14}$$

where c is a positive constant. A second specific property is that the loadings of the patterns $\mathbf{m^{(r)}}$ and $\mathbf{n^{(r)}}$, for $r = 1, \ldots, R$, are binary and of equal sign, so that each neuron is characterized by $2R$ loadings that can differ from each other only in their signs. Therefore, each neuron in a Hopfield network belongs to one of the 2^R sign combinations allowed. In terms of the

low-rank framework, Hopfield networks can therefore be described as low-rank networks with 2^R deterministic populations, which have means

$$\left(a_{m_1}^{(p)}, \cdots, a_{m_R}^{(p)}\right) = R_m \left(\pm 1, \ldots, \pm 1\right), \tag{5.15}$$

$$\left(a_{n_1}^{(p)}, \cdots, a_{n_R}^{(p)}\right) = R_n \left(\pm 1, \ldots, \pm 1\right), \tag{5.16}$$

$$\mathrm{sgn}\left(a_{m_r}^{(p)}\right) = \mathrm{sgn}\left(a_{n_r}^{(p)}\right), \tag{5.17}$$

and where there is no dispersion around the mean of each population, so that $\sigma_m^{(p)} = \sigma_n^{(p)} = 0$.

A rank-two network with four populations $P = 4$, characterized by equation 5.3 (see Figures 6A and 6C), is therefore equivalent to a two-pattern Hopfield network in the limit of no dispersion around the mean of each cluster, $\sigma_{m^2}^{(p)} = 0$. In this limit, saddle points are located at the midpoints between neighboring stable fixed points. In the more general rank-two networks in equation 5.3, where $\sigma_{m^2}^{(p)} > 0$, the saddle points between stable fixed points move farther away from the origin (such as in Figure 6B, where $\sigma_{m^2}^{(p)} = 0.3$), but the four stable fixed points remain on the vertices of a square along the axes $\kappa_1 = 0$ and $\kappa_2 = 0$. In the limit of very large $\sigma_{m^2}^{(p)}$, the saddle points between stable fixed points approach the circle that circumscribes the stable fixed points.

The rank-three network presented in equations 5.10 and 5.11 also becomes a classical Hopfield network in the limit of $\sigma_{m^2}^{(p)} \to 0$. Allowing for values $\sigma_{m^2}^{(p)} > 0$, as illustrated in Figures 6D and 6G, does not change the number of fixed points generated by the Hopfield network or their direction in collective space. These networks generate pairs of stable fixed points along the directions $\mathbf{m}^{(1)}$, $\mathbf{m}^{(2)}$, and $\mathbf{m}^{(3)}$. When $R_m R_n$ is large, additional fixed points become stable along directions $\pm \mathbf{m}^{(1)} \pm \mathbf{m}^{(2)} \pm \mathbf{m}^{(3)}$. These additional fixed points correspond to well-known spurious mixture states in Hopfield networks (Amit, Gutfreund, & Sompolinsky, 1987).

6 Approximating Dynamical Systems with Gaussian-Mixture Low-Rank Networks

In the previous section, we focused on generating multiple fixed points in an autonomous network by means of a few gaussian populations in the connectivity. More generally, as shown in section 2.3, multipopulation rank-R networks can approximate any R-dimensional dynamical system. In this section, we propose an algorithm for that purpose.

Previous works have developed algorithms for training recurrent networks to implement given dynamics that effectively used low-rank connectivity (Eliasmith & Anderson, 2003; Rivkind & Barak, 2017; Pollock &

Jazayeri, 2020). These methods rely on tuning the loadings $n_i^{(r)}$ of individual neurons, given fixed external inputs $I_i^{(s)}$ and connectivity loadings $m_i^{(r)}$. Here we focus instead on networks based on mixtures of gaussian populations, in which the couplings between individual neurons are not precisely set but instead sampled from a distribution. We extend previous methods to find the first- and second-order moments of multiple gaussian populations that approximate a given dynamical system.

Our goal is to approximate the R-dimensional dynamics specified by a vector field $G(\kappa)$:

$$\frac{d\kappa}{dt} = G(\kappa).$$

(6.1)

Our algorithm proceeds as follows. We first fix the number of gaussian populations in the network and the fraction of neurons included in each population, α_p. Depending on the complexity of the target dynamics and the required accuracy, a smaller or larger number of populations is required. Second, we set the mean and variance of the $\mathbf{m}^{(r)}$ vectors in each population, $a_{m_r}^{(p)}$ and $\sigma_{m_r^2}^{(p)}$, together with the mean and variance of the external input, $a_I^{(p)}$ and $\sigma_{I^2}^{(p)}$. We assign these parameters randomly according to a certain probability distribution. Finally, we determine the statistics of the $\mathbf{n}^{(r)}$ vectors, the only unknown in the network, using linear regression.

We define a number of set points $\{\kappa_k\}_{k=1...K}$ on which we impose that the effective flow in the low-rank network, given by equation 3.4, be equal to the target vector field:

$$G(\kappa_k) = -\kappa_k + \sum_{p=1}^{P} \alpha_p \left(a_n^{(p)} \left\langle \phi\left(\mu^{(p)}(\kappa_k), \Delta^{(p)}(\kappa_k)\right)\right\rangle \right.$$

$$\left. + \sigma_{nm}^{(p)} \kappa_k \left\langle \phi'\left(\mu^{(p)}(\kappa_k), \Delta^{(p)}(\kappa_k)\right)\right\rangle \right).$$

(6.2)

These $k = 1, \ldots, K$ set points should be relevant points of the vector field $G(\kappa)$; they can be fixed points, but can also be chosen within a grid in collective space or based on sampled trajectories of the target system (see equation 6.1). For simplicity, in equation 6.2, we are considering that the input pattern $\mathbf{I}$ is orthogonal to the connectivity patterns $\mathbf{n}^{(r)}$. It is straightforward to extend the algorithm to account for nonzero values of the parameters $\sigma_{n_r I}$.

Note that $\mu^{(p)}$ and $\Delta^{(p)}$ depend on the statistics of patterns $\mathbf{I}$ and $\mathbf{m}^{(r)}$ that are fixed (see equation 3.6), but not on $a_{n_r}^{(p)}$ and $\sigma_{m_r n_r}^{(p)}$ which we aim to determine. Equation 6.2 can therefore be written as a linear system of the form

$$\mathbf{G} = \mathbf{W}^T \mathbf{X},$$

(6.3)

where, for each individual set point, G is a vector of length R, $G = G(\kappa_k) + \kappa_k$, the vector

$$X := \left[a_{n_1}^{(1)}, \ldots a_{n_R}^{(1)}, \sigma_{m_1 n_1}^{(1)}, \ldots \sigma_{m_1 n_R}^{(1)}, \ldots \sigma_{m_R n_1}^{(1)}, \ldots \sigma_{m_R n_R}^{(1)}, \ldots a_{n_1}^{(P)} \ldots \sigma_{m_R n_R}^{(P)} \right]$$

$$(6.4)$$

has length $R(R+1)P$, and the corresponding matrix W of size $R(R+1)P \times R$. For the K set points κ_k on which we want to approximate the dynamics, we concatenate the vector G and matrix W of each point, so that they will be of size $R \cdot K$ and $R \cdot K \times (R(R+1)P)$, respectively.

The unknown values of vector $\mathbf{X}$ can now be obtained by standard linear regression as

$$X = \left(WW^T \right)^{-1} WG. \tag{6.5}$$

Often it is convenient to regularize the regression algorithm to avoid the entries of X being exceedingly large, at the cost of increasing the error in the approximation of the dynamics. Solutions with very large values of X are less robust because they produce stronger finite-size effects when sampling from the found mixture of gaussians, potentially affecting the stability of the solution. One standard possibility among many is to use ridge regression to find the unknown values,

$$X = \left(WW^T + \beta^2 I \right)^{-1} WG, \tag{6.6}$$

where β is the ridge parameter that controls the amount of regularization.

The number of populations, the level of regularization, together with the distributions chosen to fix the mean and covariance values $a_{m_r}^{(p)}, \sigma_{m_r^2}^{(p)}, a_I^{(p)}$, and $\sigma_{I^2}^{(p)}$ are hyperparameters of the algorithm. These hyperparameters can be tuned progressively by running several iterations of the algorithm. For example, a possible goal is to search for the minimal number of populations required for approximating a given dynamical system within some accuracy limits. In general, we observe empirically that the distribution underlying the fixed parameters does not play a crucial role in the accuracy of the algorithm as long as they span a wide range of values.

To illustrate the algorithm, we use a rank-two network to approximate a Van der Pol oscillator, a two-dimensional nonlinear dynamical system that generates nonharmonic oscillations. It is defined as

$$\frac{dx}{dt} = y, \tag{6.7}$$

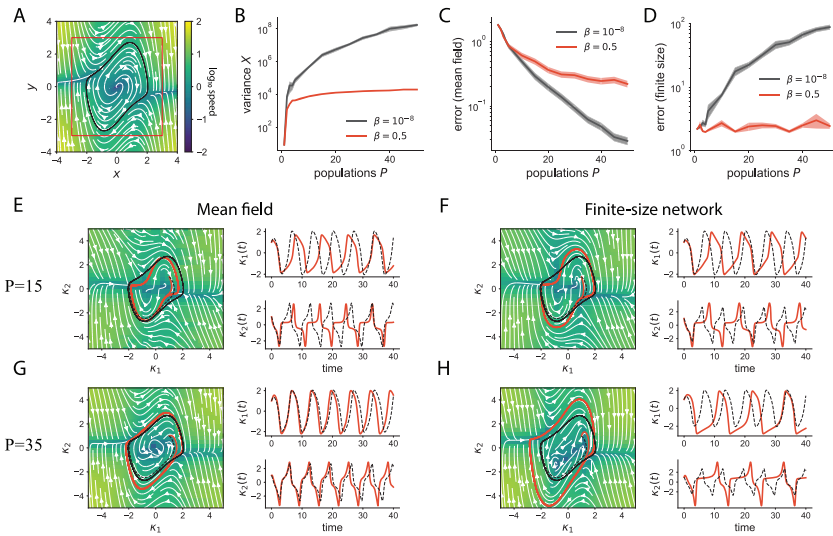$$\frac{dy}{dt} = \mu \left(1 - x^2 \right) y - x, \tag{6.8}$$

Figure 7: Approximation of a Van der Pol oscillation with multiple-population low-rank networks. (A) Dynamics of a Van der Pol oscillator ($\mu = 1$). The red square indicates the boundaries of the grid used to approximate the dynamical system. (B) Variance of the obtained parameters X for two different levels of regularization β (see equation 6.6). The variance grows extremely fast as the number of populations P is increased in nonregularized networks. Error bars represent the standard error of the mean (SEM) with 50 realizations. (C) Average approximation error by the mean-field description as a function of the number of populations. The error is measured as the average deviation from the target vector field, assessed at the grid points used for training the algorithm. (D) Average approximation error by a finite-size network, with $N = 2000$ units per population. The finite-size error quickly grows with no regularization. (E) Mean-field approximation with $P = 15$ populations. Left: dynamical landscape. The red curve corresponds to a trajectory initiated at state $(1, 1)$. In black, the corresponding trajectory of the Van der Pol oscillator. Right: trajectories as a function of time. (F) Finite-size network corresponding to the mean-field solution found in panel E. (G, H) Similar to panels E and F, with a larger number of populations, $P = 35$. The mean-field solution approximates better the Van der Pol oscillators. However, the finite-size approximation does not improve with the number of units per population kept constant ($N = 2000$). Parameters: Solutions in E–H are regularized, $\beta = 0.5$. Values $\sigma_{m_\tau^2}^{(p)}$ and $\sigma_{I^2}^{(p)}$ are initially drawn from an exponential distribution with unit variance, $a_I^{(p)}$ and $a_{m_1}^{(p)}$ are drawn from a uniform distribution with bounds -2 and 2, and $a_{m_2}^{(p)} = 0$.

where μ is a scalar parameter that controls the strength of the nonlinearity. For this example, we set $\mu = 1$ (see Figure 7A). Second, we determine the statistics of the left connectivity patterns and the external input by drawing

random values for the mean values in each population $a_I^{(p)}$ and $a_{m_r}^{(p)}$ from a zero-mean uniform distribution, and the variances $\sigma_{m_r^2}^{(p)}$ and $\sigma_{I^2}^{(p)}$ from an exponential distribution, all values of order one. As set points, we use a $K = 30 \times 30$ grid for values x and y ranging between -3 and 3 (see the red square, Figure 7A).

We first analyze the performance of the algorithm as a function of the number of populations, for two different levels of regularization. We find that regularization is necessary to avoid diverging values of the parameters in X (see Figure 7B, black versus red curve). This is due to the fact that the matrix WW^T is close to singular, so that its inverse reaches very high values. Interestingly, the error in the approximation increases only slightly with a strong level of regularization (see Figure 7C). As the number of populations is increased, the error in the approximation for all levels of regularization monotonically decreases.

The algorithm is based on the mean-field description in equation 6.2, which holds in the limit of a very large number of neurons per population. We next study whether the obtained mean-field networks describe well the dynamics in networks with a finite number of neurons per population. We therefore sampled the connectivity and input loadings from the multivariate gaussian distribution that characterizes each population. For that purpose, it is necessary to set the variances of the right connectivity patterns $\sigma_{n_r^2}^{(p)}$, which do not influence the mean-field dynamics. As a general approach, we set those variance' parameters to be as low as possible but high enough so that the correlation matrix of the multivariate gaussian distribution is positive-definite. Sampling a finite number of values from the distribution introduces deviations in the sampled mean and sampled covariance matrix that introduce additional finite-size errors in the approximated dynamics. We find that when the parameters obtained in X reach very large values, the approximation error in networks with $N = 2000$ per population is very large (see Figure 7D, black curve). However, when the algorithm is constrained by a strong level of regularization, the finite-size error remains approximately constant as the number of populations P increases (see Figure 7D, red curve).

We then show the approximated dynamics for two different numbers of populations, $P = 15$ (see Figures 7E and 7F) and $P = 35$ (see Figures 7G and 7H), in the regularized case. Fifteen populations are enough to obtain a limit cycle with similar features to the Van der Pol oscillator, although there are deviations in the shape and frequency of oscillation (see Figure 7E). The finite-size effects remain small when $N = 2000$ units are considered in each population, so that the dynamics of a simulated network resemble the mean-field approximation (see Figure 7F). For a larger number of populations, the mean-field approximation increases the accuracy considerably (see Figure 7G). However, finite-size effects increase with the number of populations, when the number of units per population is kept constant,

so that the approximation error in finite-size simulations is not necessarily reduced (see Figure 7H).

The finite-size deviations that contribute to the approximation error depend on the number of units per population N that we chose to fix $N = 2000$ in Figures 7D, 7F, and 7H. We expect a decrease in finite-size errors as N increases, so that the network dynamics resemble more closely the mean-field description. Accordingly, with a high enough number of units per population N, a finite-size network with $P = 35$ populations (as in Figure 7F) could approximate better a Van der Pol oscillator than a finite-size network with $P = 15$ populations (as in Figure 7H). A full characterization of finite-size effects is beyond the scope of this study.

This algorithm can be applied to generate any given dynamics in collective space within a finite domain. Beyond this finite domain sampled through the chosen set points, if the target vector field does not follow the required asymptotic behavior (see equation 3.14), as is the case for the Van der Pol oscillator, the network will not extrapolate to the target dynamics (region outside square $|\kappa_1, \kappa_2| > 3$ in Figures 7E to 7H, left). However, in practice, it may produce qualitatively similar dynamics: in the example of the Van der Pol oscillator, if the network is initialized at a point outside the limit cycle, the resulting trajectories still converge to the limit cycle.

7 Discussion

In this letter, we have examined the dynamics in gaussian-mixture low-rank recurrent neural networks, a class of models in which the connectivity is defined by a low-rank matrix, with connectivity patterns consisting of several populations with distinct gaussian statistics. In these networks, the collective dynamics can be described by $R + N_{in}$ collective variables, where R is the rank of the connectivity matrix and N_{in} the dimensionality of the input patterns. These collective variables form a dynamical system, the evolution of which is determined by the connectivity statistics of the populations forming the network. The rank of the network and the population structure therefore play different roles: the rank of the network sets the internal dimensionality of the dynamics and defines the corresponding collective variables, while individual populations shape the dynamics of these collective variables but do not contribute new ones. We specifically showed that in the limit of a large number of populations, this class of network displays a universal approximation property and can therefore implement a large range of dynamical systems. Having a small number of populations instead imposes constraints and limits the achievable range of dynamics.

We focused here on a specific family of distributions for the connectivity patterns, mixtures of multivariate gaussians. This choice was motivated by several considerations. First, this family of distributions can be used to approximate any multivariate distribution for the pattern loadings. Second, it

leads to a particularly simple form of dynamics for the collective variables, where the time evolution is formulated in terms of a simple effective circuit (see equation 3.11). Remarkably, in this description of the dynamics, which is exact and nonlinear, the collective variables appear to interact linearly through effective couplings and effective inputs. This allows for a particularly transparent interpretation of dynamics in terms of gain modulation. Several of our results are, however, independent of the specific assumption for the type of distribution; this is in particular the case for the influence of symmetry in the connectivity on the dynamics. When a large number of populations is needed to approximate the connectivity structure, other parametric distributions may be more suitable, and the interpretation in terms of discrete populations may not be appropriate.

Previous work (Mastrogiuseppe & Ostojic, 2018, 2019; Schuessler, Dubreuil et al., 2020) has studied the generation of low-dimensional dynamics by recurrent neural networks with a low-rank component in the connectivity, focusing on fixed points and limit cycles. In these studies, the connectivity matrix contained a random full-rank term and a low-rank term, where the low-rank loadings are randomly drawn from a single gaussian population. The full-rank term in the connectivity does not allow for the expression of the emerging dynamics directly as a low-dimensional dynamical system, as we have presented in this work (see section 3), and the study of the potential constraints on the low-dimensional dynamics.

Low-rank networks with arbitrary pattern distributions form a rich and versatile framework that encompasses a number of previously studied types of recurrent neural networks. As shown in the last part of the results, Hopfield networks storing $R \ll N$ patterns can be seen as a particular limit of gaussian-mixture low-rank networks, in which pattern loadings are binary and exhibit a specific type of symmetry. The neural engineering framework (Eliasmith & Anderson, 2003) and the manifold embedding approach (Pollock & Jazayeri, 2020) provide algorithms that implement specific low-dimensional dynamics by controlling the structure of fixed points and Jacobians using linear-regression methods. These algorithms generate recurrent networks with low-rank connectivity, in which the individual couplings between individual neurons are set to specific values. In contrast, we focus on low-rank networks in which individual couplings are sampled from an underlying distribution, and our algorithm determines the statistics of this distribution rather than individual couplings. This approach automatically endows our networks with strong robustness with respect to the deletion of individual neurons.

The dynamics of recurrent networks of rate units that include several populations have been studied in recent years (Aljadeff, Stern, & Sharpee, 2015; Aljadeff, Renfrew, Vegué, & Sharpee, 2016; Kadmon & Sompolinsky, 2015). In these works, the connectivity consisted of full-rank, random matrices in which populations define the pair-wise connectivity statistics. For example, in a two-population model of excitatory and inhibitory neurons,

the inhibitory population is defined by the connectivity statistics (mean and variance of synaptic strengths) between inhibitory neurons and the connectivity statistics toward excitatory neurons. In contrast, here we focus on low-rank structure in the connectivity matrices and specify populations in terms of the statistics for the low-rank structure vectors within each population. This allows us to go beyond random connectivity and study the effects of the rank and the number of populations independently. Directly relating the two descriptions of connectivity is a topic of ongoing work.

Our framework is also closely related to echo-state (Jaeger, 2001) and FORCE networks (Sussillo & Abbott, 2009), which rely on randomly connected recurrent networks controlled by feedback loops. Each feedback loop is mathematically equivalent to adding a unit-rank component to the connectivity matrix. Echo-state and FORCE networks therefore correspond to low-rank networks with an additionnal full-rank, random term in the connectivity (Mastrogiuseppe & Ostojic, 2018, 2019). Because the feedback loops are trained to produce specific outputs, the low-rank part of the connectivity is typically correlated to the random connectivity term (but see Mastrogiuseppe & Ostojic, 2019). Such correlations increase the dimensionality and the range of the dynamics (Schuessler, Dubreuil et al., 2020; Logiaco, Abbott, & Escola, 2019), although the low-rank connectivity structure and the number of populations still generate strong constraints. For instance, for rank-one networks with a random term in the connectivity but consisting of a single population, the fixed points are restricted to lie on a one-dimensional but nonlinear manifold, and typically at most two nontrivial stable fixed points can be generated (Schuessler, Dubreuil et al., 2020). More generally, random components in the connectivity can strongly influence learning dynamics during training (Schuessler, Mastrogiuseppe, Dubreuil, Ostojic, & Barak, 2020).

Gaussian-mixture low-rank networks, the neural engineering framework, and echo-state networks all exhibit universal approximation properties (Eliasmith, 2005; Maass, Natschläger, & Markram, 2002). It is, however, important to distinguish between several variants of this property. In our case, in analogy with the NEF, we started from an R-dimensional dynamical system fully specified by its flow function and showed that gaussian-mixture low-rank networks can approximate this flow function, provided a large number of populations is available and the flow function satisfied specific constraints. Echo-state and FORCE networks instead start by specifying a target readout, and universal approximation means that any such readout can be generated by training the feedback (Maass et al., 2007). This readout corresponds to a low-dimensional projection of a large dynamical system, and echo-state networks are free to implement any dynamical system consistent with the specified output projection. This is a major distinction with our approach and that of the NEF, where the overall dynamical system is more tightly constrained.

In this work, we have examined only networks with fixed inputs. Varying the inputs instead modifies the low-dimensional dynamics, an effect that can be understood through modulations of effective couplings that govern the interactions between collective variables. In a companion paper (Dubreuil et al., 2020), we have used gaussian-mixture, low-rank RNNs to reverse-engineer networks trained on a range of neuroscience tasks and found that gain modulation through input control underlies complex computations, such as flexible input-output mappings (Fusi, Miller, & Rigotti, 2016). Varying inputs while keeping connectivity fixed therefore has the potential of implementing a large range of dynamical systems and computations (Pollock & Jazayeri, 2020), but the full capacity of this mechanism still remains to be fully elucidated.

Appendix A: Dynamics in Multipopulation Networks

In this appendix, we derive the equation for the dynamics of a multipopulation, low-rank network, equation 3.4. We consider a low-rank network that consists of P populations, where each population is defined by different statistics of the probability distribution $P_{(p)}(\underline{m}, \underline{n}, I)$. We assume that the external input is constant in time and uncorrelated with the left connectivity patterns at the level of each population. Each neuron in the network is assigned to a population according to the probability α_p. In the following, we set the statistics of each population to be drawn from a multivariate gaussian with mean vector $\boldsymbol{a}^{(p)}$, as defined in equation 2.5 and covariance matrix $\Sigma^{(p)}$ (see equation 2.6).

The recurrent dynamics in a low-rank network are determined by equation 3.3: it consists of a sum over the N units in the network. In the limit of large networks with defined statistics, by means of the law of large numbers, this sum over N i.i.d. elements corresponds to the empirical average over the distribution of its elements. Therefore, we can replace the sum over network units for $i = 1, \ldots, N$ of loadings $\left\{n_i^{(r)}\right\}, \left\{m_i^{(r)}\right\}$, and I_i, by an integral over their probability distribution $P(\underline{m}, \underline{n}, I)$. Using this probability distribution, the recurrent dynamics in equation 3.3 can be expressed as

$$\kappa_r^{rec} = \sum_{p=1}^{P} \alpha_p \int d\underline{m} \, d\underline{n} \, dI P_{(p)}(\underline{m}, \underline{n}, I) \, n_r^{(p)} \phi\left(I^{(p)}\kappa_I + \sum_{l=1}^{R} m_l^{(p)}\kappa_l\right). \quad \text{(A.1)}$$

Note that we refer to the input loadings I as a single gaussian variable, instead of a set of gaussian variables $\underline{I}$, because there can only be one effective input pattern when the input is constant in time. We then separate the contribution of the mean a_{n_r} and the fluctuations of n_r around its mean into two terms:

$$\kappa_r^{rec} = \sum_{p=1}^{P} \alpha_p \int dI \, d\underline{m} \, P_{(p)} (\underline{m}, I) \, a_{n_r}^{(p)} \phi \left(I^{(p)} \kappa_I + \sum_{l=1}^{R} m_l^{(p)} \kappa_l \right)$$

$$+ \sum_{p=1}^{P} \alpha_p \int dn_r \, dI \, d\underline{m} \, P_{(p)} (\underline{m}, n_r, I) \left(n_r^{(p)} - a_{n_r}^{(p)} \right) \phi \left(I^{(p)} \kappa_I + \sum_{l=1}^{R} m_l^{(p)} \kappa_l \right).$$

(A.2)

This derivation is implicitly conditioned on the values of the collective variables κ_r. Under such conditioning, the argument inside the function ϕ in equation A.2 is itself a gaussian variable. Using Stein's lemma in the second term and expressing the argument of the transfer function as a single gaussian variable, we can express the dynamics as

$$\kappa_r^{rec} = \sum_{p=1}^{P} \alpha_p a_{n_r}^{(p)} \int \mathcal{D}x \, \phi \left(a_I^{(p)} \kappa_I + \sum_{s=1}^{R} a_{m_s}^{(p)} \kappa_s + x \sqrt{\sigma_{I^2}^{(p)} \kappa_I^2 + \sum_{s'=1}^{R} \sigma_{m_{s'}^2}^{(p)} \kappa_{s'}^2} \right)$$

$$+ \sum_{p=1}^{P} \alpha_p \left(\sigma_{n_r I}^{(p)} \kappa_I + \sum_{s=1}^{R} \sigma_{n_r m_s}^{(p)} \kappa_s \right)$$

$$\times \int \mathcal{D}x \, \phi' \left(a_I^{(p)} \kappa_I + \sum_{s=1}^{R} a_{m_s}^{(p)} \kappa_s + x \sqrt{\sigma_{I^2}^{(p)} \kappa_I^2 + \sum_{s'=1}^{R} \sigma_{m_{s'}^2}^{(p)} \kappa_{s'}^2} \right), \quad \text{(A.3)}$$

where $\mathcal{D}x = dx \, (2\pi)^{-\frac{1}{2}} e^{-\frac{x^2}{2}}$. Finally, using the gaussian integral notation in equation 3.8, we retrieve equation 3.5.

Appendix B: Universal Approximation of Low-Dimensional Dynamics

The universal approximation theorem for artificial neural networks (Hornik, Stinchcombe, & White, 1989; Funahashi, 1989; Cybenko, 1989) states that any piecewise-continuous bounded function $G(x)$, where x is a d-dimensional vector, can be approximated to arbitrary precision by a finite linear combination of nonlinear units having the same transfer function but different gain and thresholds. More precisely, it is possible to build an approximation $\hat{G}(x)$ of $G(x)$,

$$\hat{G}(x) = \sum_{i=1}^{N} v_i \phi \left(w_i^T x + b_i \right),$$

(B.1)

with finite integer N, and real values for $v_i \in \mathcal{R}^{d'}$, $w_i \in \mathcal{R}^d$ and $b_i \in \mathcal{R}$, so that $|G(x) - \hat{G}(x)| < \epsilon$, for any $\epsilon > 0$, given mild assumptions on the nonlinear

activation function $\phi(x)$. Historically, the universality property has been shown using a wide range of transfer functions: squashing or sigmoidal functions (Funahashi, 1989; Hornik et al., 1989; Cybenko, 1989), Heaviside functions, sinusoidal functions (Gallant & White, 1988), and radial basis functions (Park & Sandberg, 1991). Years later, it was shown that the necessary and sufficient condition on the transfer function for the universal approximation property is that $\phi(x)$ be a nonpolynomial function (Leshno et al., 1993).

There is a direct mapping between the second term of equation B.1 and the recurrent dynamics of a low-rank RNN. The recurrent dynamics in equation 3.3 can be directly mapped to equation B.1: the variables $\frac{1}{N}n_i$ correspond to v_i, m_i to w_i, and $\kappa_I I_i$ to b_i. This implies that the recurrent dynamics can approximate any flow function within a finite domain. The parameters $\frac{1}{N}n_i$, m_i, and $\kappa_I I_i$ can be adjusted independently. In particular, $\kappa_I I_i$ is independent of all the other collective variables κ_r, although the opposite is not true: the collective variables κ_r obviously depend on the external tonic input $\kappa_I I_i$. We are assuming that the temporal profile of the background input $u(t)$ is constant, so that $\kappa_I(t)$ is also constant after discarding a possible initial transient (see equation 3.2).

The dynamics of low-rank networks with multiple gaussian populations can also be mapped to the universal approximation theorem. The mean term contribution to the dynamics in equation 3.4 reads

$$\sum_{p=1}^{P} \alpha_p \boldsymbol{a_n}^{(p)} \left\langle \phi\left(\boldsymbol{a_m}^{T}\boldsymbol{\kappa} + a_I^{(p)}\kappa_I, \sigma_{I^2}^{(p)}\kappa_I^2 + \boldsymbol{\kappa}^T\sigma_{m^2}^{(p)}\boldsymbol{\kappa}\right)\right\rangle, \tag{B.2}$$

so that $\alpha_p\boldsymbol{a_n}^{(p)}$ maps to v_i, $\boldsymbol{a_m}^{(p)}$ maps to w_i, and $a_I^{(p)}\kappa_I$ is mapped to the bias term b_i. The transfer function is, however, different. In equation B.1, the nonlinear function used is $\phi(x)$, while in equation B.2, the nonlinear function used is $\langle\phi(x, \Delta(x))\rangle$. Both functions are nonlinear and nonpolynomial, so that the theorem applies in each case. The contribution given by the disorder in the population loadings, $\sigma_{m^2}^{(p)}$ and $\sigma_{I^2}^{(p)}$, are not required for the universal approximation. However, quadratic terms like the one introduced by the variance of loadings improve the approximation in terms of expressibility and efficiency (Fan, Xiong, & Wang, 2020). Overall, this means that a low-rank network with a finite number of populations can approximate any dynamical system within a bounded domain.

Appendix C: Linear Stability Matrix at Fixed Points in Networks with a Single Population

The linear dynamics of small perturbations around the fixed point $\boldsymbol{\kappa_0}$ (defined in equations 4.2) read

$$\tau\frac{d\boldsymbol{\kappa}}{dt} = -\boldsymbol{\kappa} + \left[\nabla\left(\langle\phi'\left(0, \boldsymbol{\kappa}^T\boldsymbol{\kappa}\right)\rangle\sigma_{mn}\boldsymbol{\kappa}\right)\right]_{\boldsymbol{\kappa}=\boldsymbol{\kappa_0}}\boldsymbol{\kappa}, \tag{C.1}$$

where ∇ is the vector differential operator. We apply the property $\nabla (f(\kappa) A\kappa)) = f(\kappa) A + A\kappa (\nabla f(\kappa))^T$, based on the chain rule, where A is an $R \times R$ matrix, to obtain

$$\tau \frac{d\kappa}{dt} = -\kappa + \left[\langle \phi' (0, \kappa^T \kappa) \rangle \sigma_{mn} + \sigma_{mn} \kappa \langle \nabla \phi' (0, \kappa^T \kappa) \rangle^T \right]_{\kappa = \kappa_0} \kappa. \qquad \text{(C.2)}$$

We then calculate the gradient of the gain factor. To do so, we first write explicitly the gaussian integral

$$\langle \nabla \phi' (0, \kappa^T \kappa) \rangle = \int \mathcal{D}x \nabla \phi' \left(\sqrt{\kappa^T \kappa} x \right), \qquad \text{(C.3)}$$

where $\mathcal{D}x$ is the differential element of a normally distributed variable. When the chain rule applied,

$$\langle \nabla \phi' (0, \kappa^T \kappa) \rangle = \int \mathcal{D}x \phi'' \left(\sqrt{\kappa^T \kappa} x \right) \nabla \left(x \sqrt{\kappa^T \kappa} \right)$$

$$= \int \mathcal{D}x \phi'' \left(\sqrt{\kappa^T \kappa} x \right) x \frac{\kappa}{\sqrt{\kappa^T \kappa}}. \qquad \text{(C.4)}$$

When Stein's lemma is used, the gradient of the gain factor reads:

$$\langle \nabla \phi' (0, \kappa^T \kappa) \rangle = \int \mathcal{D}x \phi'' \left(\sqrt{\kappa^T \kappa} x \right) x \frac{\kappa}{\sqrt{\kappa^T \kappa}}$$

$$= \int \mathcal{D}x \phi''' \left(\sqrt{\kappa^T \kappa} x \right) \kappa = \langle \phi''' (0, \kappa^T \kappa) \rangle \kappa. \qquad \text{(C.5)}$$

Finally, when introducing equation C.5 into equation C.2, and using the fact that $\sigma_{mn}\kappa_0 = \lambda_r \kappa_0$, the dynamics of small perturbation around the fixed point read

$$\tau \frac{d\kappa}{dt} = \left[-I + \langle \phi' (0, \kappa_0^T \kappa_0) \rangle \sigma_{mn} + \langle \phi''' (0, \kappa_0^T \kappa_0) \rangle \sigma_{mn} \kappa_0 \kappa_0^T \right] \kappa, \qquad \text{(C.6)}$$

which leads to the linear stability matrix given by 4.4.

It is important to analyze the behavior of the function $\langle \phi''' (0, \Delta) \rangle$ to assess the stability. In the limit $\Delta = 0$, the gaussian integral reduces to the evaluation of the function at zero. For a transfer function $\phi(x) = \tanh(x)$, we obtain

$$\lim_{\Delta \to 0} \langle \phi''' (0, \Delta) \rangle = \phi''' (0) = -2. \qquad \text{(C.7)}$$

In the limit of infinite Δ, the gaussian integral can be expressed as

$$\lim_{\Delta \to \infty} \langle \phi'''(0, \Delta) \rangle = \int_{-\infty}^{+\infty} dx\, \phi'''(x) = 0. \tag{C.8}$$

Furthermore, it can be shown analytically that $\langle \phi'''(0, \Delta) \rangle$ is never zero for any finite value Δ, by studying the minima of its primitive function. The primitive function of $\langle \phi'''(0, \Delta) \rangle$ is proportional to $\langle \phi'(0, \Delta) \rangle$. The primitive function has no local minima because it is bounded between zero and one, that is, a monotonically decreasing function of Δ. We can show that the function is monotonically decreasing because for sigmoidal transfer functions, for any fixed value $x > 0$, $\phi'\left(\sqrt{\Delta_1}x\right) < \phi'\left(\sqrt{\Delta_2}x\right)$ if and only if $\Delta_1 > \Delta_2$. Thus, this property is still conserved when calculating the gaussian average: $\langle \phi'(0, \Delta_1) \rangle < \langle \phi'(0, \Delta_2) \rangle$ if and only if $\Delta_1 > \Delta_2$.

Putting these analyses together, we conclude that the function $\langle \phi'''(0, \Delta) \rangle$ is -2 for $\Delta = 0$, is always smaller than zero and tends asymptotically to this upper bound as Δ approaches infinity. This result is used to study whether the linearized dynamics around fixed points in low-rank networks with a single gaussian population are stable (see equation 4.5).

Appendix D: Stability Analysis of Rank-Two Networks with Nonnormal Covariance

In section 4, we analyzed the dynamics generated by a rank-two network with one single gaussian population when the covariance matrix σ_{mn} is normal, that is, the eigenvectors are mutually orthogonal to each other. We extend here the analysis to nonnormal matrices and show that the main features of the dynamics are conserved when the correlation matrix is nonnormal.

We first studied the case of normal matrices with real eigenvalues (see Figures 3A and 3D and equations 4.2 to 4.5). The analysis showed that each real eigenvector $\mathbf{u_r}$ of the covariance matrix σ_{mn} leads to a pair of fixed points when the associated eigenvalue λ_r is larger than one.

The linear dynamics around fixed points is given by the Jacobian S_r in equation 4.4, where $r = 1, \ldots, R$. The eigenvectors of S_r coincide with the eigenvectors of σ_{mn} if the covariance matrix has real eigenvectors mutually orthogonal to each other. Using this property, we can then determine the eigenvalues of the linearized dynamics around each fixed point in equation 4.5. As a conclusion, this analysis showed that all fixed points are saddle points, except for the two fixed points in the direction of the eigenvector with the largest associated eigenvalue, where the fixed points are stable.

We now extend this analysis to networks where the covariance matrix σ_{mn} is a nonnormal matrix with real eigenvalues, in other words, a matrix with nonzero correlations between eigenvectors. Each real eigenvector $\mathbf{u_r}$

with eigenvalue $\lambda_r > 1$ generates two fixed points in the direction it spans, exactly the same way as in matrices with orthogonal eigenvectors. The fixed points are also located at the same radial distance along each eigenvector (see equation 4.3), compared to the case of normal matrices. The Jacobian matrix at the fixed point is still given by equation 4.4:

$$S_r = -I + \frac{1}{\lambda_r}\boldsymbol{\sigma_{mn}} + \langle \phi''' \left(0, \rho_r^2\right)\rangle \lambda_r \rho_r^2 \boldsymbol{u_r u_r}^T. \tag{D.1}$$

When the covariance matrix $\boldsymbol{\sigma_{mn}}$ is a nonnormal matrix, the eigenvectors of S_r are not equal to the eigenvectors of the covariance. Finding an analytical expression for the eigenvectors of S_r is in general a challenging problem. However, we can still calculate the eigenvalues of S_r and show that they remain unchanged in networks with normal covariances (see equation 4.5).

We consider the plane spanned by the eigenvector $\boldsymbol{u_r}$ and any other eigenvector of the covariance matrix $\boldsymbol{u_{r'}}$. We consider only the linearized dynamics around the fixed point along this plane. We project S_r onto this plane and indicate it with the notation $[S_r]$, which is a 2×2 block submatrix of the full Jacobian S_r:

$$[S_r] = -[I] + \frac{1}{\lambda_r}[\boldsymbol{\sigma_{mn}}] + \langle \phi''' \left(0, \rho_r^2\right)\rangle \lambda_r \rho_r^2 \boldsymbol{u_r u_r}^T. \tag{D.2}$$

The vector $\boldsymbol{u_r}$ is still an eigenvector of $[S_r]$. Its associated eigenvalue is

$$\gamma_r = \langle \phi''' \left(0, \rho_r^2\right)\rangle \lambda_r \rho_r^2 \tag{D.3}$$

which is a negative number. This implies that all fixed points are stable in the radial direction. Although calculating the second eigenvector is not obvious, we can calculate the second eigenvalue of $[S_r]$ by calculating its trace and subtracting γ_r:

$$\gamma_{r'} = \text{tr}\left([S_r]\right) - \gamma_r. \tag{D.4}$$

Using the linearity of the trace operator, the expression simplifies to

$$\gamma_{r'} = -1 + \frac{\lambda_{r'}}{\lambda_r}. \tag{D.5}$$

Therefore, the eigenvalues of the full Jacobian S_r, given by equations D.3 and D.5 coincide with the eigenvalues of the Jacobian in the case of normal matrices (see equation 4.5). The fixed points along the eigenvector direction $\boldsymbol{u_r}$ are stable if λ_r is the largest eigenvalue of $\boldsymbol{\sigma_{mn}}$. Otherwise, the fixed point is a saddle point.
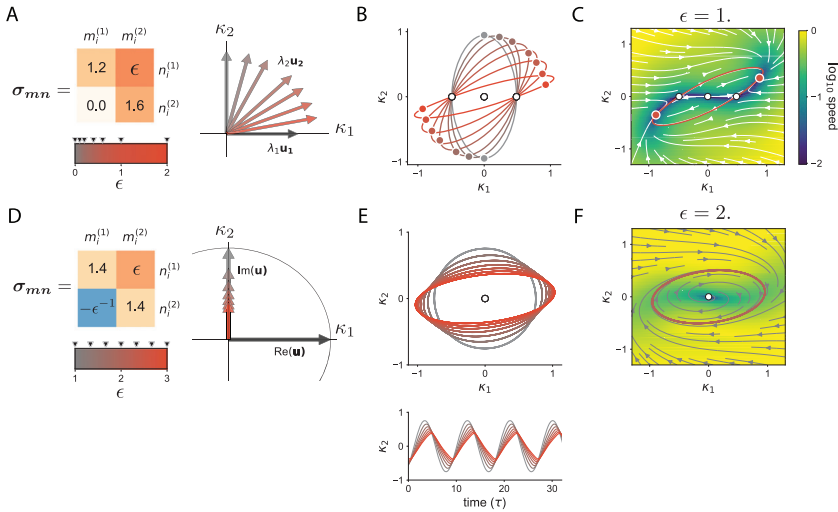
Figure 8: Dynamics in rank-two networks with nonnormal covariances σ_{mn}. (A) Covariance matrix σ_{mn} with eigenvalues $\lambda_1 = 1.2$ and $\lambda_2 = 1.6$. The free parameter ϵ controls the angle between the eigenvectors, as shown in the right. (B) Stable fixed points (colored dots) and saddle points (white dots) generated by the nonnormal covariance matrix for different values of ϵ. The colored lines indicate the trajectories between saddle points and stable fixed points. As ϵ increases, the location of the stable fixed points moves closer to the horizontal line while keeping the same radial distance. The stability of the fixed points does not change with ϵ. (C) Example of the full dynamics in collective fixed space for a fixed value of $\epsilon = 1$. (D) Covariance matrix σ_{mn} with eigenvalues $\lambda = 1.4 \pm 1$. The free parameter $\epsilon > 1$ controls the relative norm of the imaginary part of the eigenvector $\mathbf{u}$, without modifying the fixed points. (E) Top: Limit cycles that emerge in the dynamics for different values of ϵ. When the covariance matrix is normal ($\epsilon = 1$, gray line), the limit cycle has a circular shape. As ϵ increases, the limit cycle loses the circular symmetry and resembles an ellipse. Bottom: Projection of the collective variable κ_2 as a function of time. As ϵ increases, the activity loses its sinusoidal shape while keeping the same frequency. (F) Example of the dynamical landscape in collective space for a fixed value of $\epsilon = 2$.

Figures 8A and 8C show how the dynamics vary when the correlation matrix σ_{mn} is nonnormal but the eigenvalues remain constant. The parameter ϵ determines the degree of correlation between eigenvectors: as ϵ increases, the two eigenvectors become more strongly correlated (see Figure 8A). This correlation moves the direction of the associated fixed point, while keeping the radial distance constant (see Figure 8B). The eigenvalues of the linear dynamics around the fixed points do not vary. Figure 8C shows an example of the dynamical landscape in collective space. It is

interesting to note that the trajectories that go from the saddle points toward the stable fixed points resemble ellipses, similar to the case of normal σ_{mn} (see Figures 3A and 3D). However, the stable fixed points are no longer located at the farthest point of this curve from the origin.

In summary, we showed that the number of fixed points and the stability in a low-rank network with one single population depends only on the eigenvalues of $\sigma_{\mathbf{mn}}$, and not on its eigenvectors. When the eigenvectors are all mutually orthogonal, we can calculate analytically the eigenvalues and eigenvectors of the linearized dynamics around each fixed point. When the eigenvectors are not all mutually orthogonal, the eigenvectors of the linearized dynamics change, but not the eigenvalues. This suggests that increasing the correlation between eigenvectors introduces a continuous deformation of the collective space while the timescales of the linearized dynamics around fixed points are preserved.

We observed numerically that the previous result holds as well for covariance matrices σ_{mn} with complex eigenvalues. Each pair of complex conjugate eigenvalues $\lambda \pm i\omega$ with eigenvector $\mathbf{v_1} + i\mathbf{v_2}$ generates a limit cycle on the plane spanned by $\mathbf{v_1}$–h$\mathbf{v_2}$ if $\lambda > 1$. If σ_{mn} is a normal matrix, such that vectors $\mathbf{v_1}$ and $\mathbf{v_2}$ have the same norm and are orthogonal to each other, the limit cycle is a circle. In that case, we showed in section 4 that the frequency of the limit cycle is given by $\frac{\omega}{\lambda}$.

Adding correlations between the real and imaginary part of the eigenvectors $\mathbf{v_1}$ and $\mathbf{v_2}$, or changing their relative norm, will change the shape of the limit cycle to a curve resembling an ellipse. Figures 8D an 8F show an example of a covariance matrix σ_{mn} where the nonnormality is controlled by parameter ϵ without affecting the eigenvalues. As ϵ increases, the relative norm of the imaginary and real part of the eigenvector $\mathbf{u}$ changes (see Figure 8D). This correlation changes the shape of the limit cycle, from a circle when ϵ is one (normal case; see the gray line in Figure 8E) to a closed curve resembling an ellipse as ϵ increases (see Figures 8E and 8F, the gray line). The frequency of oscillation along the limit cycle, however, stays invariant as ϵ is varied (see Figure 8E bottom).

To sum up, also in the case of complex eigenvalues, the correlations between eigenvectors deform the collective space, introducing a continuous mapping between the circular limit cycle and the new limit cycle, which resembles an ellipse. However, the temporal features of the dynamical landscape, such as the frequency of the limit cycle, remain constant.

Real and complex eigenvalues are combined in networks with rank three or larger. The same stability rules hold: the dynamic structure (limit cycle or fixed point) generated by the eigenvalue with the largest real part is stable, while the other ones are not stable in all directions. Figures 9A and 9B show an example of a rank-three network, whose connectivity matrix has a real eigenvalue λ_1 and a pair of complex conjugate eigenvalues λ_2 and λ_3. The real part of all eigenvalues is larger than one, so the real eigenvalue leads to a pair of fixed points, and the complex eigenvalues generate a limit
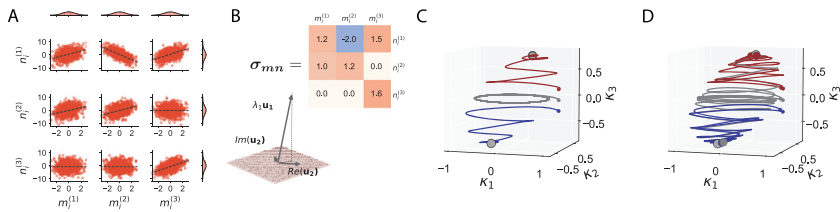
Figure 9: Dynamics in a rank-three network with a single gaussian population - connectivity matrix combining real and complex eigenvalues. (A) Scatter plot between the loadings of connectivity patterns $m_i^{(r)}$ and $n_i^{(r)}$. $\sigma = 1.6$ and $\sigma_\omega = 0.8$. (B) Covariance matrix of the singular vectors (top) and sketch of the eigenvectors (bottom). The eigenvalues are $\lambda_1 = 1.6$ and $\lambda_{2,3} = 1.2 \pm \sqrt{2}$. The real eigenvector $\mathbf{u}_1$ is not orthogonal to the plane spanned by the real and imaginary part of the complex eigenvectors $\mathbf{u}_2$. The real and imaginary parts of the complex eigenvectors span the horizontal plane (shaded in gray) and do not have the same norm. (C) Mean-field dynamics (see equation 4.1) for three trajectories starting at different initial conditions. Each color indicates a different trajectory. When the network is initialized in the horizontal plane (gray trajectory), the activity oscillates within the noncircular limit cycle. Otherwise it converges to one of the two stable fixed points, located in the direction of the eigenvector $\mathbf{u}_1$. (D) Same trajectories as in panel C, in finite-size simulations, for three different connectivity matrices. The trajectories always end up in one of the two stable fixed points, even if initialized in the horizontal plane (gray trajectories). Parameters: $N = 1000$, $\sigma_{n_r^2} = 9$.

cycle. Given that the real eigenvalue λ_1 is larger than the real part of the other eigenvalues, the fixed points are stable. The limit cycle is marginally stable in the plane spanned by the real and imaginary parts of the complex eigenvector of λ_2 but unstable in any other direction. Therefore, trajectories starting in the plane converge to the limit cycle in the mean-field equation (see the gray trajectory in Figure 9C). Small perturbations, such as those introduced by finite-size effects, make these trajectories deviate from the limit cycle and converge to one of the two stable fixed points (see the gray trajectories, Figure 9D).

Acknowledgments

Code Availability

Code and trained models are available at https://github.com/emebeiran/low-rank2020.

References

Aljadeff, J., Renfrew, D., Vegué, M., & Sharpee, T. O. (2016). Low-dimensional dynamics of structured random networks. *Physical Review E*, *93*(2), 022302.

Aljadeff, J., Stern, M., & Sharpee, T. (2015). Transition to chaos in random networks with cell-type-specific connectivity. *Physical Review Letters*, *114*(8).

Amit, D. J., Gutfreund, H., & Sompolinsky, H. (1987). Statistical mechanics of neural networks near saturation. *Annals of Physics*, *173*(1), 30–67.

Barak, O. (2017). Recurrent neural networks as versatile tools of neuroscience research. *Current Opinion in Neurobiology*, *46*, 1–6.

Buonomano, D. V., & Maass, W. (2009). State-dependent computations: Spatiotemporal processing in cortical networks. *Nature Reviews Neuroscience*, *10*(2), 113–125.

Chaisangmongkon, W., Swaminathan, S. K., Freedman, D. J., & Wang, X. J. (2017). Computing by robust transience: How the fronto-parietal network performs sequential, category-based decisions. *Neuron*, *93*(6), 1504–1517.

Churchland, M. M., & Shenoy, K. V. (2007). Temporal complexity and heterogeneity of single-neuron activity in premotor and motor cortex. *Journal of Neurophysiology*, *97*(6), 4235–4257.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems*, *2*(4), 303–314.

Doya, K. (1993). *Universality of fully-connected recurrent neural networks* (Technical Report 1, 1–6). Department of Biology, University of California, San Diego.

Dubreuil, A., Valente, A., Beiran, M., Mastrogiuseppe, F., & Ostojic, S. (2020). *Complementary roles of dimensionality and population structure in neural computations*. bioRxiv 2020.07.03.185942.

Eliasmith, C. (2005). A unified approach to building and controlling spiking attractor networks. *Neural Computation*, *17*(6), 1276–1314.

Eliasmith, C., & Anderson, C. H. (2003). *Neural engineering: Computation, representation, and dynamics in neurobiological systems*. Cambridge, MA: MIT Press.

Fan, F., Xiong, J., & Wang, G. (2020). Universal approximation with quadratic deep networks. *Neural Networks*, *124*, 383–392.

Funahashi, K.-I. (1989). On the approximate realization of continuous mappings by neural networks. *Neural Networks*, *2*(3), 183–192.

Fusi, S., Miller, E. K., & Rigotti, M. (2016). Why neurons mix: High dimensionality for higher cognition. *Current Opinion in Neurobiology*, *37*, 66–74.

Gallant, A. R., & White, H. (1988). There exists a neural network that does not make avoidable mistakes. In *Proceedings of the International Conference on Neural Networks* (pp. 657–664). Piscataway, NJ: IEEE.

Gallego, J. A., Perich, M. G., Naufel, S. N., Ethier, C., Solla, S. A., & Miller, L. E. (2018). Cortical population activity within a preserved neural manifold underlies multiple motor behaviors. *Nature Communications*, *9*(1), 1–13.

Gao, P., Ganguli, S., Battaglia, F. P., & Schnitzer, M. J. (2015). On simplicity and complexity in the brave new world of large-scale neuroscience. *Current Opinion in Neurobiology*, *32*, 148–155.

Hennequin, G., Vogels, T. P., & Gerstner, W. (2014). Optimal control of transient dynamics in balanced networks supports generation of complex movements. *Neuron*, *82*(6), 1394–1406.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. In *Proceedings of the National Academy of Sciences of the United States of America*, *79*(8), 2554–2558.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, *2*(5), 359–366.

Jaeger, H. (2001). *The "echo state" approach to analysing and training recurrent neural networks—with an erratum note 1* (GMD Report 148). Bonn: German National Research Center.

Kadmon, J., & Sompolinsky, H. (2015). Transition to chaos in random neuronal networks. *Physical Review, X*, *5*(4), 041030.

Leshno, M., Lin, V. Y., Pinkus, A., & Schocken, S. (1993). Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. *Neural Networks*, *6*(6), 861–867.

Logiaco, L., Abbott, L., & Escola, S. (2019). *A model of flexible motor sequencing through thalamic control of cortical dynamics*. bioRxiv:2019.12.17.880153.

Maass, W., Joshi, P., & Sontag, E. D. (2007). Computational aspects of feedback in neural circuits. *PLOS Comput. Biol.*, *3*(1), 165.

Maass, W., Natschläger, T., & Markram, H. (2002). Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation*, *14*(11), 2531–2560.

Machens, C. K., Romo, R., & Brody, C. D. (2010). Functional, but not anatomical, separation of "what" and "when" in prefrontal cortex. *Journal of Neuroscience*, *30*(1), 350–360.

Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, *503*(7474), 78–84.

Mastrogiuseppe, F., & Ostojic, S. (2018). Linking connectivity, dynamics, and computations in low-rank recurrent neural networks. *Neuron*, *99*(3), 609–623.

Mastrogiuseppe, F., & Ostojic, S. (2019). A geometrical analysis of global stability in trained feedback networks. *Neural Computation*, *31*(6), 1139–1182.

Nakatsukasa, Y. (2019). *The low-rank eigenvalue problem*. arXiv:1905.11490.

Park, J., & Sandberg, I. W. (1991). Universal approximation using radial-basis-function networks. *Neural Computation*, *3*(2), 246–257.

Pollock, E., & Jazayeri, M. (2020). Engineering recurrent neural networks from task-relevant manifolds and dynamics. *PLOS Comput. Biol.*, *16*(8), e1008128.

Rabinovich, M., Huerta, R., & Laurent, G. (2008). Neuroscience: Transient dynamics for neural processing. *Science*, *321*(5885), 48–50.

Rajan, K., Harvey, C. D., & Tank, D. W. (2016). recurrent network models of sequence generation and memory. *Neuron*, *90*(1), 128–142.

Remington, E. D., Egger, S. W., Narain, D., Wang, J., & Jazayeri, M. (2018). A dynamical systems perspective on flexible motor timing. *Trends in Cognitive Sciences*, *22*(10), 938–952.

Remington, E. D., Narain, D., Hosseini, E. A., & Jazayeri, M. (2018). Flexible sensori-
    motor computations through rapid reconfiguration of cortical dynamics. *Neuron*,
    98(5), 1005–1019.

Rigotti, M., Barak, O., Warden, M. R., Wang, X. J., Daw, N. D., Miller, E. K., & Fusi,
    S. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature*,
    497(7451), 585–590.

Rivkind, A., & Barak, O. (2017). Local dynamics in trained recurrent neural networks.
    Physical Review Letters, *118*(25), 258101.

Saxena, S., & Cunningham, J. P. (2019). Towards the neural population doctrine. *Cur-
    rent Opinion in Neurobiology*, *55*, 103–111.

Schuessler, F., Dubreuil, A., Mastrogiuseppe, F., Ostojic, S., & Barak, O. (2020). Dy-
    namics of random recurrent networks with correlated low-rank structure. *Physi-
    cal Review Research*, *2*(1), 013111.

Schuessler, F., Mastrogiuseppe, F., Dubreuil, A., Ostojic, S., & Barak, O. (2020). The
    interplay between randomness and structure during learning in RNNs. In H.
    Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, & H. Lin (Eds.), *Advances in
    neural information processing systems*, *33*. Red Hook, NY: Curran.

Sohn, H., Narain, D., Meirhaeghe, N., & Jazayeri, M. (2019). Bayesian computation
    through cortical latent dynamics. *Neuron*, *103*(5), 934–947.

Sussillo, D., & Abbott, L. (2009). Generating coherent patterns of activity from chaotic
    neural networks. *Neuron*, *63*(4), 544–557.

Sussillo, D., Churchland, M. M., Kaufman, M. T., & Shenoy, K. V. (2015). A neural
    network that finds a naturalistic solution for the production of muscle activity.
    Nature Neuroscience, *18*(7), 1025–1033.

Vyas, S., Golub, M. D., Sussillo, D., & Shenoy, K. V. (2020). Computation through
    neural population dynamics. *Annual Review of Neuroscience*, *43*(1), 249–275.

Wang, J., Narain, D., Hosseini, E. A., & Jazayeri, M. (2018). Flexible timing by tem-
    poral scaling of cortical responses. *Nature Neuroscience*, *21*(1), 102–112.

Yang, G. R., Joglekar, M. R., Song, H. F., Newsome, W. T., & Wang, X.-J. (2019). Task
    representations in neural networks trained to perform many cognitive tasks. *Na-
    ture Neuroscience*, *22*(2), 297–306.