Neural manifold capacity via the cavity method

George Cai*
Department of Neurobiology, Harvard Medical School
(Dated: May 16, 2025)

I provide a concise summary of the theory of neural manifold capacity, which provides a normative measure of the efficiency of neural representations for linear classification tasks. I present recent work on an analytically tractable formulation for neural manifolds with correlations. Then, I walk through an alternative derivation to the replica theory. Lastly, I conduct simulations in a synthetic model of correlated manifolds to verify the analytical capacity formula and provide intuition for the effects of correlations on the manifold capacity.

I. INTRODUCTION

Human and animal brains have a remarkable capability of invariant object recognition despite enormous natural variations in context, pose, position, lighting, occlusion, and other variables [1]. Only recently, through advances in parallel computation and large-scale neural network, artificial neural networks have achieved comparable or even super-human performance in image recognition tasks [2]. Early efforts to understand why deep neural networks succeeded in image classification led to the striking finding that visual stimulus-response filters in artificial neural networks resembled the Gabor-like stimulus-response kernels in the visual processing pathway in animal brains [3]. These observations suggest there may be unifying, normative computational principles that shape both artificial and biological neural networks that are performant in image classification. In sensory neuroscience, an early example of a normative principle was the *efficient coding hypothesis* [4]. The hypothesized objective of early sensory systems was to maximize the information content of neural activity subject to energy constraints. This notion has been formalized into a sparse encoding model in which a small set of visual neurons respond to each input stimulus due to L_1-penalty on neural activations [5]. However, the predictions of this theory do not describe or explain how more "cognitive" regions of the visual stream transform inputs for invariant object recognition.

A. Perceptron Capacity

An alternative notion of coding efficiency has emerged from theoretical study of perceptron classification of points and subsequently, "neural manifolds." The basic set-up is the following: Given P randomly placed points in $\mathbb{R}^N$ and a random labeling of the points into two classes, what is the probability there exists a separating hyperplane for the points in the two classes? Cover showed through a simple counting argument that for points in general position, the probability of linear separability exhibits a phase transition in $\alpha = P/N$ in the limit $P, N \to \infty$ (see fig. 1**A**) [6] . The critical "capacity" α_c at which linear separability vanishes is a notion of efficiency because a neural representation that permits linear separation of a larger number of randomly labeled points relative to the neurons or dimensions used is encoding stimuli more efficiently for classification tasks. In the case of randomly labeled points, $\alpha_c = 2$.

Gardner extended the capacity theory to support vector machines, which classify points with a "margin" distance κ between the nearest points and the hyperplane (see fig. 1**b**) [7]. The problem was formulated in terms of the volume of solution hyperplanes $\mathbf{w} \in \mathbb{R}^N, ||\mathbf{w}||_2^2 = N/\kappa$ that satisfy margin-classification constraints for every data point and performed a replica calculation of the logarithm of the solution volume. There is a critical value of α such that the solution volume vanishes, and the points are no longer linearly separable with margin κ.

To perform quenched averages of the log-volume, the Gardner formulation requires the data points to be drawn from either a multivariate Gaussian distribution or N-dimensional random sign distribution. However, this is rarely a good description of either artificial or biological neural network representations. A recent line of work has generalized the Gardner calculation to *neural manifolds*, which are not formal manifolds but rather collections of data points that share a label [8, 9]. A simple example is the set of images of dogs and cats as represented by a particular layer in a neural network, decpited in fig. 1**C**. Neural manifolds allow for variations in features such as pose or viewing angle. To handle the variable geometries of manifolds, the theory considers the convex hull of each manifold. By linearly separating all convex hulls, the data points within the manifolds are guaranteed to be separated as well. Then, the notion of a support vector, which is a data point with distance from the hyperplane exactly equal to the margin κ, can be generalized to *anchor points*, which can be thought of as a "representative" point for each manifold with respect to the classification problem. The manifold capacity problem has been solved for uncorrelated and correlated manifolds via the replica method [9, 10]. I state the manifold assumptions in section II A. Then, I use the cavity method to more simply and rigorously calculate the manifold capacity when manifolds have correlations in their geometric properties

* georgecai@g.harvard.edu

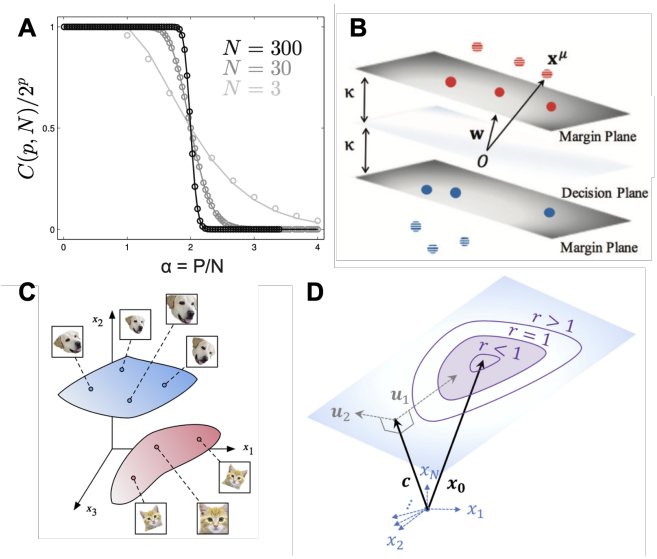

FIG. 1. **A** Cover's theorem gives the probability of linear separability given random labels as a function of the load $\alpha = P/N$. Figure from [11]. **B** Support vector machine classification of data points with margin κ. **C** Example manifolds of cat images and dog images. Variations within manifold include pose and size of the animal. **D** 2-dimensional neural manifold in affine subspace of $\mathbb{R}^N$ with center $\mathbf{x}_0$, basis vectors $\mathbf{u}_1$, $\mathbf{u}_2$, orthogonal translation vector of the subspace $\mathbf{c}$, and global manifold scale r. Panels **B**-**D** from [9].

II B. Both the cavity and replica approaches arrive at an expression that has been verified in simulations in [9, 10], and I reproduce a subset of these results in III.

B. Contribution

This project presents an alternative proof of the capacity of correlated manifolds via the cavity method by modifying an existing calculation for the case of uncorrelated manifolds [12]. It also provides simulations to validate the accuracy of the formula.

II. THEORETICAL FRAMEWORK

A. Data manifold model

Consider a dataset of P manifolds, each parameterized by coordinates $\mathcal{S} \subseteq \mathbb{R}^{D+1}$ and a binary label $y^\mu \in \{-1, +1\}$. Embed each manifold in $\mathbb{R}^N$ through random embeddings u^μ_{ia} indexed by $\mu = 1, \ldots, P$ manifolds, $i = 1, \ldots, N$ neural features, and $a = 1, \ldots, D+1$ manifold subspace axes. Here, D is the number of axes along which the manifold varies, and the $(D+1)$th element corresponds to the center of the manifold, which is schematized in fig. 1**D**. Denote the ith row of the μth embedding matrix as $\mathbf{u}^\mu_i \in \mathbb{R}^{D+1}$. We assume random embeddings of the manifolds with i.i.d. components u^μ_{ia} following:

$$\langle u^\mu_{ia} \rangle = 0, \tag{1}$$

$$\langle u^\mu_{ia} u^\nu_{jb} \rangle = \frac{1}{N} C^{\mu\nu}_{ab} \delta_{ij}. \tag{2}$$

In a more compact notation, the correlation matrix has dimensions $C \in \mathbb{R}^{P(D+1) \times P(D+1)}$, and we denote the (μ, ν)th block as $C^{\mu\nu} \in \mathbb{R}^{(D+1)\times(D+1)}$. Similarly, when we drop the manifold index μ for the embedding vector, we concatenate over the column vectors: $\mathbf{u}_i \in \mathbb{R}^{P(D+1)}$.

A point with manifold coordinates $\mathbf{s} \in \mathcal{S}^\mu \subseteq \mathbb{R}^{D+1}$ on manifold μ specifies an N-dimensional datapoint given by

$$x^\mu_i(\mathbf{s}) = \sum_{a=1}^{D+1} u^\mu_{ia} s_a = (\mathbf{u}^\mu_i)^\top \mathbf{s}, \;\; i = 1, \ldots, N. \tag{3}$$

The last coordinate s_{D+1} specifies the displacement from the origin, and the remaining coordinates specify the shape in a D-dimensional subspace orthogonal to the displacement. Because our classifier is linear, we can just consider the convex hull of $\mathcal{S}$. To classify every point correctly, we require for all $\mathbf{s} \in \mathcal{S}^\mu, \mu = 1, \ldots, P$,

$$y^\mu \mathbf{w}^\top \mathbf{x}^\mu(\mathbf{s}) \geq 1. \tag{4}$$

We introduce the compact notation that considers the projection of the solution hyperplane vector onto the manifold embeddings

$$\mathbf{V}^\mu = y^\mu \mathbf{w}^\top \mathbf{u}^\mu, \;\; \mathbf{V}^\mu \in \mathbb{R}^{D+1} \tag{5}$$

so the classification condition eq. (4) becomes

$$(\mathbf{V}^\mu)^\top \mathbf{s} \geq 1. \tag{6}$$

The support function of $\mathcal{S}$

$$g_{\mathcal{S}^\mu}(\mathbf{V}^\mu) = \min_{\mathbf{s} \in \mathcal{S}} (\mathbf{V}^\mu)^\top \mathbf{s} \tag{7}$$

returns the smallest projection of $\mathcal{S}^\mu$ onto $\mathbf{V}^\mu$. Now, we define an anchor point for the μth manifold as

$$\mathbf{a}^\mu = \nabla_{\mathbf{s}} g_{\mathcal{S}^\mu}(\mathbf{V}^\mu) \tag{8}$$

$$= \arg\min_{\mathbf{s} \in \mathcal{S}^\mu} (\mathbf{V}^\mu)^\top \mathbf{s}; \;\; \mathbf{a}^\mu \in \mathbb{R}^{D+1}, \tag{9}$$

where the gradient equals the arg min because $\mathcal{S}^\mu$ are convex hulls. We are now ready to introduce the convex optimization problem for the maximum margin hyperplane $\mathbf{w}$, which has Lagrangian

$$\mathcal{L}(\mathbf{w}, \{\lambda^\mu\}_{\mu=1}^P) = \frac{1}{2} \mathbf{w}^\top \mathbf{w} - \sum_{\mu=1}^P \lambda^\mu (g_{\mathcal{S}^\mu}(\mathbf{V}^\mu) - 1) \tag{10}$$

The KKT conditions are given by:

$$w_i = \sum_{\mu=1}^P \lambda^\mu y^\mu (\mathbf{u}^\mu_i)^\top \mathbf{a}^\mu \tag{11}$$

$$g_{\mathcal{S}^\mu}(\mathbf{V}^\mu) - 1 \geq 0 \tag{12}$$

$$\lambda^\mu \geq 0 \tag{13}$$

$$\lambda^\mu (g_{\mathcal{S}^\mu}(\mathbf{V}^\mu) - 1) = 0. \tag{14}$$

Without loss of generality, assume $y^\mu = +1$ since the data and labels are sampled independently and $(y^\mu)^2 = 1$. We are now ready to set up the bipartite system for analysis in the next section.

B. Bipartite system

Consider the bipartite system

$$\mathbf{V}^\mu = \sum_{i=1}^{N} w_i \mathbf{u}_i^\mu + \mathbf{I}^\mu \tag{15}$$

$$w_i = \sum_{\mu=1}^{P} \lambda^\mu y^\mu (\mathbf{u}_i^\mu)^\top \mathbf{a}^\mu + I_i, \tag{16}$$

where I_i and $\mathbf{I}^\mu$ are infinitesimal source terms used in computing response functions.

C. Feature cavity

We introduce a "feature cavity" w_0 with weak coupling u_{0a}^μ to the "datum variables" (vector-Lagrange multipliers) $\{\lambda^\mu \mathbf{a}^\mu\}_{\mu=1}^P$. The perturbation is

$$\delta(\lambda^\mu \mathbf{a}^\mu) = \sum_{\nu=1}^{P} \frac{d(\lambda^\mu \mathbf{a}^\mu)}{d\mathbf{I}^\nu} \mathbf{u}_0^\nu w_0. \tag{17}$$

Propagating the perturbation from the datum variables to the feature cavity, we have

$$w_0 = \sum_{\mu=1}^{P} (\mathbf{u}_0^\mu)^\top (\lambda^\mu \mathbf{a}^\mu + \delta(\lambda^\mu \mathbf{a}^\mu)) + I_0 \tag{18}$$

$$= \sum_{\mu=1}^{P} (\mathbf{u}_0^\mu)^\top \left(\lambda^\mu \mathbf{a}^\mu + \sum_{\nu=1}^{P} \frac{d(\lambda^\mu \mathbf{a}^\mu)}{d\mathbf{I}^\nu} \mathbf{u}_0^\nu w_0 \right) + I_0. \tag{19}$$

We define local field and self-coupling terms

$$\eta_0 = \sum_{\mu=1}^{P} (\mathbf{u}_0^\mu)^\top (\lambda^\mu \mathbf{a}^\mu) \tag{20}$$

$$F_{00} = \sum_{\mu,\nu=1}^{P} (\mathbf{u}_0^\mu)^\top \frac{d(\lambda^\mu \mathbf{a}^\mu)}{d\mathbf{I}^\nu} \mathbf{u}_0^\nu. \tag{21}$$

Then, the single site equation is

$$w_0 = \eta_0 + F_{00} w_0 + I_0. \tag{22}$$

The cavity construction ensures $\mathbf{u}_0^\mu$, $\lambda^\mu \mathbf{a}^\mu$ and $\frac{d(\lambda^\mu \mathbf{a}^\mu)}{d\mathbf{I}^\nu}$ are all independent of each other in this expression.

Next, we calculate the disorder-averaged statistics. By the Central Limit Theorem, η_0 is Gaussian since the correlations $\langle u_{0a}^\mu u_{0b}^\nu \rangle$ decay as $N \to \infty$. Then,

$$\langle \eta_0 \rangle = 0 \tag{23}$$

$$\langle \eta_0^2 \rangle = \langle (\vec{\lambda \mathbf{a}})^\top \mathbf{u}_0 \mathbf{u}_0^\top (\vec{\lambda \mathbf{a}}) \rangle \tag{24}$$

$$= \frac{1}{N} \langle (\vec{\lambda \mathbf{a}})^\top C (\vec{\lambda \mathbf{a}}) \rangle \tag{25}$$

$$= \frac{1}{N} \langle ||C^{1/2} (\vec{\lambda \mathbf{a}})||_2^2 \rangle, \tag{26}$$

where we write the concatenation of the vector Lagrange-multipliers as $\vec{\lambda \mathbf{a}} \in \mathbb{R}^{P(D+1)}$. We assume the self-coupling F_{00} is self-averaging and compute the mean:

$$\langle F_{00} \rangle = \langle \mathbf{u}_0^\top \chi \mathbf{u}_0 \rangle \tag{27}$$

$$= \langle \mathrm{Tr}[\mathbf{u}_0^\top \chi \mathbf{u}_0] \rangle \tag{28}$$

$$= \langle \mathrm{Tr}[\chi \mathbf{u}_0 \mathbf{u}_0^\top] \rangle \tag{29}$$

$$= \mathrm{Tr}[\chi C], \tag{30}$$

where $\chi \in \mathbb{R}^{P(D+1) \times P(D+1)}$ and the blocks are $\chi^{\mu\nu} = \frac{d(\lambda^\mu \mathbf{a}^\mu)}{d\mathbf{I}^\nu}$. Then, dropping the source term, we have a single-site self-consistent equation:

$$w_0 = \frac{\eta_0}{1 - F_{00}} \tag{31}$$

$$= S^W \eta_0, \tag{32}$$

with weight-response function $S^W \equiv \left\langle \frac{dw_0}{dI_0} \right\rangle = \frac{1}{1 - \langle F_{00} \rangle}$. Imposing the weight norm $||w||_2^2 = N/\kappa^2$, we have

$$\kappa^{-2} = (S^W)^2 \langle \eta_0^2 \rangle \tag{33}$$

$$= (S^W)^2 \frac{P}{P} \frac{1}{N} \langle ||C^{1/2} (\vec{\lambda \mathbf{a}})||_2^2 \rangle \tag{34}$$

$$\implies \alpha^{-1}(\kappa) = \frac{1}{P} \kappa^2 (S^W)^2 \langle ||C^{1/2} (\vec{\lambda \mathbf{a}})||_2^2 \rangle. \tag{35}$$

This completes the feature cavity calculation for capacity, and we will see that the datum cavity calculation in the next section enables us to eliminate the weight-response function and replace the norm over vector-Lagrange multipliers with a quadratic program.

D. Datum cavity

Now we introduce the datum cavity $(\lambda^0 \mathbf{a}^0)$ with weak couplings $\mathbf{u}_i^0$. The small perturbation to the existing weights is

$$\delta w_i = \sum_{j=1}^{N} \frac{dw_i}{dI_j} (\mathbf{u}_j^0)^\top (\lambda^0 \mathbf{a}^0). \tag{36}$$

The projection of the solution weight vector onto the manifold basis, $\mathbf{V}^0$, after whitening by $(C^{00})^{-1/2} \in$

$\mathbb{R}^{(D+1)\times(D+1)}$, becomes

$$(C^{00})^{-1/2}\mathbf{V}^0 = \sum_{i=1}^{N} w_i (C^{00})^{-1/2}\mathbf{u}_i^0 + \tag{37}$$

$$\sum_{i,j=1}^{N} \frac{dw_i}{dI_j}(C^{00})^{-1/2}\mathbf{u}_i^0(\mathbf{u}_j^0)^\top(\lambda^0\mathbf{a}^0) + (C^{00})^{-1/2}\mathbf{I}^0 \tag{38}$$

$$= \mathbf{T}^0 + \mathbf{F}^{00}(\lambda^0\mathbf{a}^0) + (C^{00})^{-1/2}\mathbf{I}^0, \tag{39}$$

with $\mathbf{T}^0 \equiv \sum_{i=1}^N w_i (C^{00})^{-1/2}\mathbf{u}_i^0$, $\mathbf{F}^{00} \equiv \sum_{i,j=1}^N \frac{dw_i}{dI_j}(C^{00})^{-1/2}\mathbf{u}_i^0(\mathbf{u}_j^0)^\top$. By the Central Limit Theorem, $\mathbf{T}^0$ is Gaussian with

$$\langle \mathbf{T}^0 \rangle = 0 \tag{40}$$

$$\langle \mathbf{T}^0(\mathbf{T}^0)^\top \rangle = \left\langle \sum_{i,j=1}^{N} w_i w_j (C^{00})^{-1/2}(\mathbf{u}_i^0)(\mathbf{u}_j^0)^\top((C^{00})^{-1/2})^\top \right\rangle \tag{41}$$

$$= \frac{1}{N}\sum_{i=1}^{N}\langle w_i^2\rangle \tag{42}$$

$$= \sum_{i,j=1}^{N} w_i w_j \frac{1}{N}(C^{00})^{-1/2}(C^{00}\delta_{ij})(C^{00})^{-1/2} \tag{43}$$

$$= \frac{1}{\kappa^2}. \tag{44}$$

We assume the self-coupling matrix is order one and self-averaging with mean

$$\langle \mathbf{F}^{00} \rangle = \sum_{i,j=1}^{N} \left\langle \frac{dw_i}{dI_j} \right\rangle (C^{00})^{-1/2}\frac{1}{N}C^{00}\delta_{ij} \tag{45}$$

$$= \left\langle \frac{dw_0}{dI_0} \right\rangle (C^{00})^{1/2} \tag{46}$$

Together, the single-site equation after dropping the source term is

$$(C^{00})^{-1/2}\mathbf{V}^0 = \mathbf{T}^0 + S^W(C^{00})^{1/2}(\lambda^0\mathbf{a}^0). \tag{47}$$

Combining eq. (47) with the remaining KKT conditions eqs. (12) to (14), we have KKT conditions for the convex optimization problem

$$\mathbf{V}^0 = \arg\min_{\mathbf{V}\in\mathbb{R}^{(D+1)}} \frac{1}{2}||\mathbf{T}^0 - (C^{00})^{-1/2}\mathbf{V}||_2^2, \tag{48}$$

subject to $g_{\mathcal{S}}(\mathbf{V}) \geq 1$. We can rearrange eq. (47) into

$$(S^W)^2(\lambda^0\mathbf{a}^0)^T C^{00}(\lambda^0\mathbf{a}^0) = \tag{49}$$

$$\min_{\mathbf{V}, g_{\mathcal{S}}(\mathbf{V})\geq 1} ||\mathbf{T}^0 - (C^{00})^{-1/2}\mathbf{V}||_2^2 \tag{50}$$

$$(S^W)^2\left\langle ||(C^{00})^{1/2}(\lambda^0\mathbf{a}^0)||_2^2 \right\rangle = \tag{51}$$

$$\left\langle \min_{\mathbf{V}, g_{\mathcal{S}}(\mathbf{V})\geq 1} ||\mathbf{T}^0 - (C^{00})^{-1/2}\mathbf{V}||_2^2 \right\rangle. \tag{52}$$

Now, we extend eq. (52) to the unrolled vectors $V, \mathbf{T}, \vec{\lambda\mathbf{a}} \in \mathbb{R}^{P(D+1)}$ since the datum cavity calculation is symmetric for any index μ, and we substitute into the right hand side of eq. (35):

$$\alpha^{-1}(\kappa) = \frac{\kappa^2}{P}\left\langle \min_{V, g_{\mathcal{S}}(V^\mu)\geq 1, \forall\mu} ||\mathbf{T} - C^{-1/2}V||_2^2 \right\rangle_{\mathbf{T}} \tag{53}$$

where $V \in \mathbb{R}^{P(D+1)}$ and $\mathbf{T} \sim \mathcal{N}(0, \kappa^{-2}I_{P(D+1)})$. This completes the cavity method calculation for the capacity of correlated manifolds. The analytic expression requires finding the anchor points $\mathbf{a}^\mu$ for each sampled $\mathbf{T}$, and then solving the quadratic program with affine constraints. In the next section, we show that the formula is accurate in the toy case of Gaussian point clouds.

III. TOY MODEL SIMULATION

We simulate Gaussian point cloud "manifolds" with the following correlation structure:

$$C_{ab}^{\mu\nu} = \begin{cases} [(1-\rho_a^2)\delta^{\mu\nu} + \rho_a^2]\delta_{ab}; & a \neq D+1 \\ [(1-\rho_c^2)\delta^{\mu\nu} + \rho_c^2]\delta_{ab}; & a = D+1, \end{cases} \tag{54}$$

where ρ_c and ρ_a are the correlations of the centroids and axes of the point clouds respectively. The ground truth manifold capacity is simulated as

$$\alpha_{sim} = \frac{P}{N^*} \tag{55}$$

where $1 \leq N^* \leq N$ is the neural dimension such that a random projection of the data manifolds into $\mathbb{R}^{N^*}$ has a 50% chance of being linearly separable. This definition is equivalent to finding the load $\alpha = P/N$ such that 50% of random labelings of the manifolds are linearly separable.

In fig. 2**A**,**C**, we find that the analytic formula for the capacity provides a good approximation to the ground-truth simulated capacity as we vary the center correlation and axis correlation respectively. In fig. 2**B** we find that increasing the correlations of the manifold centers decreases capacity. At the limit of perfectly correlated manifold centers, capacity would be 0 because the manifolds would be impossible to linearly separate. At best, completely uncorrelated manifold centers will have the highest linear separability. In fig. 2**D**, we find that increasing the inter-manifold axis correlation increases capacity. Intuitively, if all manifold axes are perfectly correlated, then it is easier to linearly separate the manifolds because they vary or extend in the same directions. By contrast, if the manifolds are all randomly oriented, then they have a greater tendency to extend in a way that decreases the margin.

IV. CONCLUSION

In this project, I provide an exposition on classification capacity, a measure of the efficiency of a neural representation for linear classification. I overview an extension

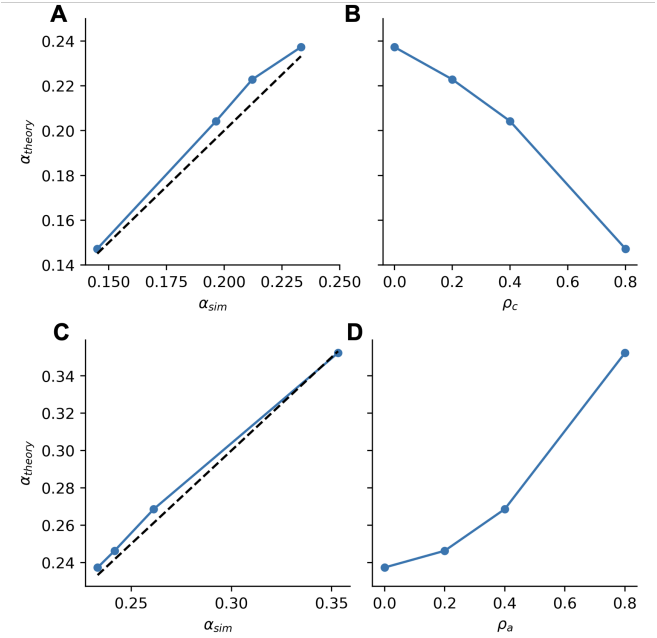

FIG. 2. Correlated Gaussian point cloud simulation. **A** Simulated and analytical capacity as center correlation is varied. **B** Theoretical capacity as a function of inter-manifold center correlation. **C** Simulated and analytical capacity as axis correlation is varied. **D** Theoretical capacity as a function of inter-manifold axis correlation.

of Gardner's replica theory approach to neural manifolds with and without correlations, and I provide a simpler alternative proof of the correlated manifold capacity using the Cavity method. Lastly, I provided simulations on synthetic data to show the accuracy of the derived formula and illustrate the opposite effects of manifold center and axis correlations on the capacity.

Although it is omitted in this project, the development of the manifold capacity theory has enabled geometric interpretations of neural representations in both artificial and biological neural networks. These developments represent a novel and promising approach to the analysis of high-dimensional neural data that describes the data geometry without dimensionality reduction or visualization techniques.

ACKNOWLEDGMENTS

I gratefully acknowledge the support of Will Slatton, Chi-Ning Chou, and SueYeon Chung in discussions of the theory and illustrative examples.

[1] H. S. Seung and D. D. Lee, Cognition. the manifold ways of perception, Science **290**, 2268 (2000).
[2] K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition, arXiv [cs.CV] (2015).
[3] C. Olah, N. Cammarata, L. Schubert, G. Goh, M. Petrov, and S. Carter, enAn overview of early vision in InceptionV1, Distill **5**, e00024.002 (2020).
[4] H. B. Barlow, *Sensory Communication*, edited by W. A. Rosenblith, The MIT Press (MIT Press, London, England, 2012).
[5] B. A. Olshausen and D. J. Field, Emergence of simple-cell receptive field properties by learning a sparse code for natural images, Nature **381**, 607 (1996).
[6] T. M. Cover, Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition, IEEE Trans. Electron. Comput. **EC-14**, 326 (1965).
[7] E. Gardner and B. Derrida, Optimal storage properties of neural network models, J. Phys. A Math. Gen. **21**, 271 (1988).
[8] S. Chung, D. D. Lee, and H. Sompolinsky, Linear readout of object manifolds, Phys. Rev. E. **93**, 060301 (2016).
[9] S. Chung, D. D. Lee, and H. Sompolinsky, Classification and geometry of general perceptual manifolds, Phys. Rev. X **8**, 031003 (2018).
[10] A. J. Wakhloo, T. J. Sussman, and S. Chung, Linear classification of neural manifolds with correlated variability, Phys. Rev. Lett. **131**, 027301 (2023).
[11] E. Orhan, Cover's function counting theorem (1965).
[12] D. G. Clark and H. Sompolinsky, Simplified derivations for high-dimensional convex learning problems, arXiv [cond-mat.dis-nn] (2024).
[13] D. H. Hubel and T. N. Wiesel, Receptive fields and functional architecture of monkey striate cortex, J. Physiol. **195**, 215 (1968).