NORTHWESTERN UNIVERSITY

# Recurrent inhibition within olfactory networks shapes noise correlations and stimulus discrimination

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Applied Mathematics

By

Angelia Wang

EVANSTON, ILLINOIS

December 2022

# ABSTRACT

Recurrent inhibition within olfactory networks shapes noise correlations and stimulus discrimination

Angelia Wang

Each second, living organisms take in sensory input from an ever-changing environment and respond appropriately. Identifying and contextualizing stimuli is critical for survival, and it often necessitates distinguishing between sensory experiences that are similar to each other. Pattern separation characterizes the mechanisms by which neuronal networks extract and highlight differences between similar stimulus-induced activity patterns, making it easier for higher-order brain centers to distinguish between them.

Within the olfactory system, pattern separation is made possible by its extensive circuitry through which neurons communicate and interpret each other's activities. Previous work discusses how specific shared connectivity can induce pairwise correlations between neurons and how correlations reduce the total information a network can carry, but the roles of specific recurrent connectivities and mechanistic interactions between cells within the bulb are not well understood. In this thesis, we investigate how different motifs of recurrent connectivity shape noise correlations, how variances and pairwise covariances

are impacted by stimulus shape and measurement duration, and how all these quantities impact performance in a two-class stimulus discrimination problem.

The olfactory bulb contains recurrently-connected mitral cells (MCs) and granule cells (GCs), whose excitatory-inhibitory interactions generate network oscillations in the gamma frequency range. In Chapter 2, we find that recurrent inhibition from the GCs induces a gamma rhythm characterized by alternating volleys of MC and GC spikes and that MC fluctuate on the scale of the gamma period. On short timescales, MCs fire synchronously (i.e. within the same volley of spikes); over longer timescales, they reduce each other's activity via shared inhibition. Correspondingly, pairwise covariance between two MCs is positive for short measurement durations and negative in measurements larger than the period of the gamma rhythm.

In Chapter 3, we quantify network performance using linear discriminability, which is governed by the average difference in response ($\Delta\vec{\mu}$) between two stimuli and the sum of response covariances ($\Sigma$). Linear discriminability measures the dissimilarity between two stimulus-evoked patterns and may represent a cortical neuron's readout of MC activity. We find that inhibition worsens stimulus discrimination in a network comprising independent pairs of MC-GC reciprocal connections ("single connections"), but that this reduction is largely dominated by average difference in response $\Delta\vec{\mu}$. Inhibition also reduces variance in these single connected networks, which benefits discriminability and partially offsets the effect of $\Delta\vec{\mu}$. Conversely, in a network with all-to-all coupling ("global connections"), small amounts of inhibition improve discriminability despite reducing $\Delta\vec{\mu}$, and noise correlations in $\Sigma$ become increasingly beneficial as the stimuli become more similar.

We assess both optimal and exploratory linear discriminability, which constitute different ways a network can perform discrimination. Optimal discriminability $\mathcal{F}_{opt}$ represents the best-case-scenario performance of a network that optimally weights each MC input in order to maximize stimulus separability, and random (exploratory) discriminability $\mathcal{F}_{random}$ represents the performance of a network that has not learned how to discriminate an odor pair and weights MC inputs randomly. Networks connected with either single connections or global connections perform similarly in both metrics – inhibition delivered through single connections worsens discriminability and small inhibition delivered through global connections improves discriminability. However, a key difference between the two metrics is that $\mathcal{F}_{random}$ accounts for noise in dimensions (MC inputs) which have the same average activity across both stimuli. As a result, preferentially inhibiting these distracting cells reduces the total noise in the system, which improves $\mathcal{F}_{random}$.

In Chapter 4, motivated by the idea that cortical responses may result from a few spikes soon after stimulus onset, we implement a sniff cycle and take measurement windows sequentially along the sniff. We find that discriminability in the steady-state regime matches discriminability over the later portion of a sniff, but that there may also be substantial improvements during early inhalation. Specifically inhibition delivered through either single-connected and all-to-all networks improves both $\mathcal{F}_{opt}$ and $\mathcal{F}_{random}$, but only at the specific times when MCs with the largest differences between stimuli are active, whereas inhibition that preferentially suppresses strongly-spiking uninformative cells produces strong sustained improvement in discriminability over the entire inhalation. We

also find that inhibition that reduces overall discriminability in the steady-state may produce alternating periods of improved and worsened discriminability, even if the overall time-averaged discriminability is reduced.

Chapters 5 and 6 include model implementation and discussion of this thesis work, respectively.

In Chapter 7, I describe my CAR-T modeling project completed as a Clinical Pharmacology summer intern at Takeda Pharmaceuticals. While not directly related to the primary thesis work, this research nevertheless showcases an example of using mathematical modeling to further understanding of biological mechanisms. In this project, I selected a case study of CD19-targeting CAR-T used to treat both immunodeficient mice and a cohort of clinical cancer patients, and I constructed a physiology-based mechanistic model that simultaneously captured data from both cohorts. In doing so, the model predicted key mechanistic differences in how preclinical (non-human) species and clinical patients respond to this specific CAR-T construct. Specifically, T-cells expand approximately four times more rapidly in clinical patients compared to immunodeficient mice, and it also has approximately five-fold higher tumor-killing efficiency after infusion. This work may inform clinical dosing strategies in the future and provides a framework for other modeling efforts in translational medicine.

The appendix includes more detailed analyses of some of the methods we use, as well as various exploratory analyses that informed model development.

## Acknowledgements

It takes a village to graduate a PhD student, and I want to specifically recognize some of the very important people (and animals) who were integral in this journey.

First, I would like to recognize my advisor, Professor Hermann Riecke, for the thousands of hours of support he has provided during these last several years. It is very rare to land an advisor who is willing to wade through the metaphorical muck with you, schedule mealtimes and naptimes around your meetings, and tirelessly put up with your general incompetency for the sake of molding you into a marginally better scientist. I am very grateful. Thank you also to the Riecke family – Marge, Charlie, Toby, Sasha, and the kitties – for welcoming me into your home.

I would also like to thank Professors Bill Kath and Danny Abrams for being on my committee and providing guidance through this journey, and making time to sit in my annual reviews, read this thesis, and generally see me through this process despite their busy schedules and new additions (!) to their families. Their feedback has been infinitely valuable, and I learned a lot through presenting to them.

The remaining acknowledgements comprise the very many individuals without whom I would not have gotten here.

To Lila, the hairy gremlin who lives in my house and screams very loud.

To my initial Northwestern cohort: Lionel, Colton, Adam, Biyi, Joel, Can, and Alex. To the late nights kvetching over asymptotics homework, to sneaking in libations past

the professors, to fighting over the office thermostat until it broke. People say that grad school is a special time, and I owe it to you guys for making the initial transition especially memorable.

To the various ESAM members I had the great honor of befriending and commiserating with. Brennan, Brit, and Steph, for getting me into running and staying in touch after graduating. Avi, Hongyu, Xize, Lae Un, Bennet, for the humor, the homemade bread rolls, for sharing your families' cooking and being a fun research family. Andre, the best pandemic office neighbor with great music taste. Kaitlin and Becca (and Brit, as of January 2023!), for demonstrating that women can have a PhD and also a baby. This is by no means an exhaustive list.

To the various talented, hospitable, generous friends who made my grad school experience indelibly better: Yue, Karl, Sophia, Nicole, Jason, Jerome, Josh, Callen, Tiffany, Luke, Shannon, Rachel, Jackie, Jace, and many many others. To Matt Rivers and Kurt and Northwestern BLAST, for providing an amazing dance community and welcoming me in. To Gimli, Thayer, Stevie, Harley, and Gina, for being adorable.

To my college friends: Lindy, Kevin, Lizzie, Yifan, Natasha, Jazzy, Brown Ballroom. Thank you for cheerleading, for your Christmas cards, for sharing pictures of your animals.

To the lovely humans I've had the pleasure of living with: Sarah, Jari, and Michael. Thanks for applesauce cake, for Love Island UK and Tiger King, and for introducing me to heavy euro board games. For not being alarmed when you found me still working at @#$% in the morning.

To Lin-Manuel Miranda and the team behind the Encanto soundtrack, for getting me through Chicago winter.

To my talented and hilarious internship cohort: Keyur, Ahmed, Erfan, Jordan. I'm so excited to meet you all in person and see what amazing scientists you all become.

To the various mentors over the years who've given me indispensable advice: Helen, Peter B., Ganesh, Aman, Weirong, Viva, Tao, Jin, Shalla. Thanks for taking a chance on a clueless applied mathematics student and coaching me through some truly epic projects.

To my parents, who are not quite getting the doctor daughter they envisioned. ☺

To John, for helping me become a better version of myself.

To Chau ā yí, for always being kind and teaching me how to make my favorite foods.

To Bridget, who has worked so hard to forge her own path in life and deserves every bit of success she has found. For inspiring me to live loudly and unabashedly.

To Peter, for being the most incredible companion these last few years. For the endless encouragement and support, and for being someone with whom I can be my ugliest and most vulnerable self. For doing the unquestionable majority of cooking, cleaning, cat-poop-scooping, hairball-wiping, unglamorous chores around the house while I was interviewing and writing this dang thesis. For accompanying me into our next adventure.

From the bottom of my heart, thank you.

## List of abbreviations

| Name | Unit/Value | Meaning |
|---|---|---|
| OB | | Olfactory bulb |
| OSN | | Olfactory sensory neuron |
| MC | | Mitral cell |
| GC | | Granule cell |
| $S^{(i)}$ | $kHz$ | Poisson rate of Stimulus $(i)$ |
| $\Delta\vec{\mu}$ | $kHz_{Nm\times 1}$ | average difference in response rates |
| $\Sigma$ | $kHz^2_{Nm\times Nm}$ | Sum of covariance matrices of rates |
| $\mid M(f) \mid^2$ | | Power spectral density |
| $\xi,\ \zeta$ | | $\delta-$correlated Gaussian noise terms |
| $\mathcal{F}_{opt}$ | | Optimal Fisher linear discriminability |
| $\mathcal{F}_{uncorr}$ | | Uncorrelated linear discriminability |
| $\mathcal{F}_{random}$ | | Random linear discriminability |
| $w$ | | Inhibitory coupling |
| $T$ | ms | Measurement duration |
| $W_{gm}$ | | Weight matrix from GCs to MCs[†] |
| $W_{mg}$ | | Weight matrix from MCs to GCs |
| $W_{om}$ | | Weight matrix from OSNs to MCs |
| $C$ | 281 pF | membrane capacitance |
| $gL$ | 30 nS | leak conductance |
| $EL$ | -70 mV | resting potential |
| $\Delta T$ | 2 mV | exponential slope factor |
| $V_T$ | -50.4 mV | resting membrane threshold |
| $V$ | mV | Voltage |
| $\bar{x}$ | mV | Low-pass filtered voltage |
| $\tau$ | mS | Synaptic time constant |

Table 0.1. Parameters and abbreviations

[†] Conventionally, matrix notation $W_{ij}$ indicates input from source $j$ to recipient $i$. The author found this indexing more intuitive.

# Table of Contents

# List of Figures

CHAPTER 1

# Introduction

The mammalian brain is a powerful processing center that translates raw sensory input from the environment into behavioral decisions, often in seconds or faster. This ability to accurately identify environmental information with speed and precision is crucial for many survival-related tasks, including identifying food, responding to predators, and socializing with other members of the species. Our sense of smell in particular, which came under unprecedented global popularity in Spring 2020, is one of the oldest sensory modalities and helps us understand and perceive the world.

With hundreds of different chemical receptor types, the human olfactory system can distinguish (estimatedly) thousands of odorants [**1, 2, 3**], and many animals rely on olfaction as their primary sensory modality for day-to-day tasks [**4**]. The discriminatory ability of the olfactory system is largely due to the complex, extensive circuitry through which its networks of neurons communicate with each other. Understanding the dynamics of these circuits can yield insights towards how animals effectively discriminate between different stimuli.

Early processing of olfactory stimuli is composed of the following steps. Upon inhalation, volatile chemical compounds in the air move into the nasal cavity and land on olfactory receptor neurons. These neurons respond by sending electrical signals downstream to the olfactory bulb (OB), where principal neurons called mitral cells (MCs) respond and produce an activity representation specific to that odor. Inhibitory neurons

called granule cells (GCs) are reciprocally connected to MCs through dendrodendritic connections [**5**]. GCs process MC activity, dynamically reshaping it before MC activity is read out downstream by higher-order cortical cells responsible for decision making. During this intermediate reformatting step, the neuronal network extracts and highlights differences from similar patterns of stimulus-induced neural activity; this process is known as pattern separation and is important for stimulus discrimination [**6, 7, 8**].

Many experimental studies highlight the importance of GCs in enhancing discrimination tasks [**9, 10, 11**], but specific relevant mechanisms and roles of recurrently-connected neurons within the bulb are not currently well understood. In this thesis, we discuss specifically how recurrent connectivity within the OB shapes noise correlations and pairwise covariances between MCs, and how these resulting noise correlations impact stimulus discriminability.

Previous efforts studying various sensory systems begin to explain how shared connections and correlations impact discriminability. Computational studies of the visual system, in which neurons interact with each other through local excitation paired with far-reaching inhibition (commonly called "Mexican hat" connectivity), find that pairwise correlations between neurons reduce the total amount of information that the network can carry, suggesting that they are detrimental for discriminability [**12, 13, 14**]. Other studies find that whether or not correlations reduce information is largely determined by the neurons' tuning curve [**15, 16**], specific forms of their correlation matrices [**17, 18**], or the circuit motifs (feedforward-recurrent) connecting them [**19**].

Correlations that limit information limitations can nevertheless improve behavioral stimulus discrimination [**20**] or otherwise improve encoding precision [**21**]. While apparently paradoxical, this is based off the fact that discrimination is not solely determined by total available information but rather relies on different functional modalities. For example, the OB engages in fast network oscillations, along which MCs tend to spike synchronously in response to shared connections [**22**]. GCs induce synchronicity on these short timescales, resulting in positive correlations between pairs of MCs that reduce the total information but enhance propagation [**23, 24**]. In larger measurement windows however, cells become anticorrelated due to the recurrent nature of the network and indirect competition between MCs. Thus, there exists a tradeoff in stimulus discrimination; increased synchronicity at shorter time scales increases propagation to downstream brain regions, whereas pairwise anticorrelations at larger time scales increases the amount of total information available in the system [**23**].

We contribute to these efforts using a framework specific to olfactory system functionality, and we supplement our discussion of noise correlations with a broader analysis of odor discriminability and the various factors contributing to it: network connectivity, stimulus profile, measurement duration, and steady-state vs. aligned-to-sniff. Our framework consists of a network of exponential integrate-and-fire neurons that models the core excitatory-inhibitory interaction between MCs and GCs in the OB (Figure 1.1a), and we assess how these networks perform in a two-class discrimination problem. Here, MC activity in response to an odorant comprises a noisy high-dimensional cloud representation, and the distinction or "discriminability" of two similar odorants depends both on the average distance between two "stimulus clouds" and the spread of each one (Figure 1.1b). Taken

together, this thesis explores how different motifs of recurrent connectivity influence the magnitudes and signs of variance and pairwise correlations, the roles of stimulus shape and magnitude in shaping qualitative features of noise, the importance of measurement duration and placement, and how noise correlations impact stimulus discriminability.

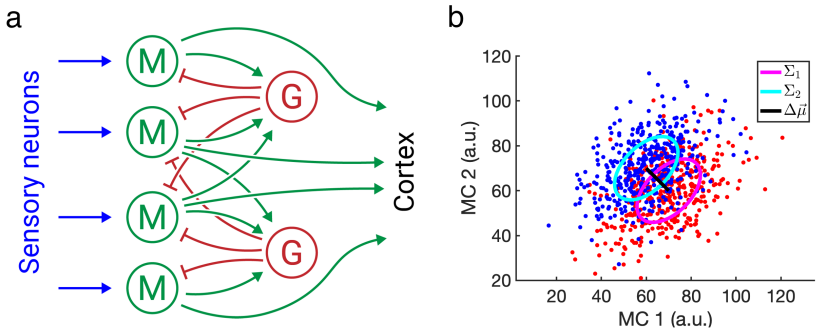


Figure 1.1. Network schematic, visual representation of discriminability metric. a) Schematic of olfactory bulb. Mitral cells (green) excite granule cells (red) and are reciprocally inhibited by them. Mitral cells are stimulated by Olfactory Sensory Neurons (blue) and also synapse onto neurons within piriform cortex. b) Schematic of linear discriminant analysis in two dimensions. Red and blue clouds represent noisy responses to two stimuli; solid black line is the vector of differences of the means. Teal and magenta ellipses represent covariance matrices of the two stimuli.

CHAPTER 2

# Noise correlations are shaped by network connectivity, stimulus shape, and measurement duration.

## 2.1. Single reciprocal connection

Our investigation of how recurrent inhibition shapes noise correlations begins with the simplest connected network: a single MC-GC reciprocal pair (Figure 2.1a).

The MC and GC are each modeled as an exponential integrate-and-fire neuron with no refractory period (Equation 2.1). When its voltage reaches a threshold $V_T$, we record a spike and reset the voltage to resting potential $E_L$. These neurons have an FI curve that appears linear, especially at larger inputs (Figure 2.1b). Consistent with experimental observations [**25**], their variability also increases with input (Figure 2.1b).

$$C\frac{dv}{dt} = -g_L\left(v - E_L\right) + g_L\Delta_T e^{\frac{v-v_T}{\Delta_T}} + I \tag{2.1}$$

$$v > V_T \rightarrow v = E_L.$$

Here, $C$ is the membrane capacitance, $g_L$ is the leak conductance, $E_L$ is the resting potential, $\Delta_T$ is a slope factor, and $I$ represents input. Parameters and general structural model are borrowed from [**26**] and listed in Table 0.1.

The mitral cell receives input from an "olfactory sensory neuron" (OSN), represented as a Poisson spike train whose rate is fixed in time. Whenever the OSN spikes, it increments an exponentially-decaying low-pass filter variable.

$$\tau_{\mathrm{OSN}} \frac{d}{dt} \bar{x}_{\mathrm{OSN}} = -\bar{x}_{\mathrm{OSN}} + \sum_{\text{OSN spike } j} \delta(t - t_j) \tag{2.2}$$

Similarly, when the MC or GC reaches threshold, it spikes and increment its respective low-pass-filter variable. The excitatory coupling strength from the MC to GC is fixed, and the inhibitory coupling strength from the GC to the MC is parametrized with a scaling variable $w$, which ranges from 0 to 50 (arbitrary units). Here and in all subsequent analyses in the main body of the thesis, we use the following values of $w$:

$$w = 0,\ 0.1,\ 0.2,\ 0.4,\ 0.6,\ 0.8,\ 1,\ 2,\ 5,\ 10,\ 20,\ 50. \tag{2.3}$$

These numbers are chosen arbitrarily but cover ranges of $w$ at which key qualitative results occur.

$$\begin{aligned}
\tau_{\mathrm{MC}} \frac{d}{dt} \bar{x}_{\mathrm{MC}} &= -\bar{x}_{\mathrm{MC}} + \sum_{\text{MC spike } k} \delta(t - t_k) \\
\tau_{\mathrm{GC}} \frac{d}{dt} \bar{x}_{\mathrm{GC}} &= -\bar{x}_{\mathrm{GC}} + \sum_{\text{GC spike } i} \delta(t - t_i) \\
I_{\text{to MC}} &= -w \times W_{gm} \times \bar{x}_{\mathrm{GC}} + W_{om} \times \bar{x}_{\mathrm{OSN}} \\
I_{\text{to GC}} &= W_{mg} \times \bar{x}_{\mathrm{MC}},
\end{aligned} \tag{2.4}$$

We first evolve the reciprocally connected pair for 50,000 ms using a standard Forward Euler algorithm with $dt = 0.02$ ms, and we keep track of MC spike history. To estimate the finite-time variance of rates, we take randomly placed measurements along the spike history, count the spikes occurring in each measurement, and calculate the variance of rates from those samples. Measurement durations vary from $T = 2$ to $T = 100$ ms.

Note that since the time series is finite and measurements are taken randomly throughout the time series, samples are not necessarily independent from each other. However, estimations of variance do not seem to be spuriously affected at our choices of measurement size nor bin size (see Appendix A.2).

Inhibition reduces MC firing rate, and MCs with higher baseline activities are suppressed more strongly than those with lower activities (Figure 2.1c). The latter follows from the a combination of the frequency-input relationship and network reciprocity; the FI curve is linear (concave up at small inputs near threshold), and active mitral cells more strongly drive their granule cells. However, MC firing rate decreases slower than on the order $\frac{1}{1+w}$ (Figure 2.1c, inset), which would be the expected scaling if both MCs and GCs' firing rates were linear in their inputs.

Variance of rates also trends lower with inhibitory strength (Figure 2.1d), especially at small $w$. We plot our neuron's variance against the corresponding variances of an artificial Poisson spike train with the same mean rate (Figure 2.1d, dashed lines), in order to establish baseline comparisons against known variance scaling.

Note: Figures 2.1c and 2.1d use a logarithmic horizontal scale in order to better show behavior at small $w$. Horizontal variable is $1 + w$ in order to include $w = 0$. Horizontal axis in the inset of figure 2.1c is on a linear axis to show that MC firing rate does not

scale linearly with $(1+w)$, which would be the expected scaling if both MC and GC firing rates were linear in their inputs.

At all input levels and at small inhibition, the variance of a reciprocally-inhibited MC reduces more steeply than in its corresponding Poisson spike train. Due to network reciprocity, mitral cells receive feedback inhibition that directly results from their own activity, which strongly reduces their noise.

At larger input levels, variance does not decrease monotonically with $w$; while unintuitive, this results from the size of the measurement window relative to the period of the system rhythm (Figure 2.2).

We next assess how variance of rates changes with measurement duration $T$. As in our analysis of mean rates, variance is compared against that of a corresponding artificial Poisson spike train of the same mean rate, whose variance decreases linearly with $T$ (Appendix A.1). To better highlight qualitative features and account for this known scaling, we multiply variance of rates by the size of the measurement duration. At small inhibitions, scaled variance decreases monotonically with $T$ (Figure 2.2a). However, at sufficiently strong inhibitions and large inputs, scaled variance nonmonotonically changes with $T$. There is notably a rise at small window sizes before a dip, followed by subsequent smaller rises and dips (Figure 2.2bc). Moreover, peaks move towards larger window sizes with increasing inhibition, consistent with the previous observation that variance at a fixed window size may increase with larger $w$ (Figure 2.1d).

Nonmonotonic behavior at large inhibitions is caused by an emergent oscillation within the network and depends on the inhibitory strength and the speed of granule cell response (see Appendix A.3). Whereas small amounts of inhibition do not drive any apparent

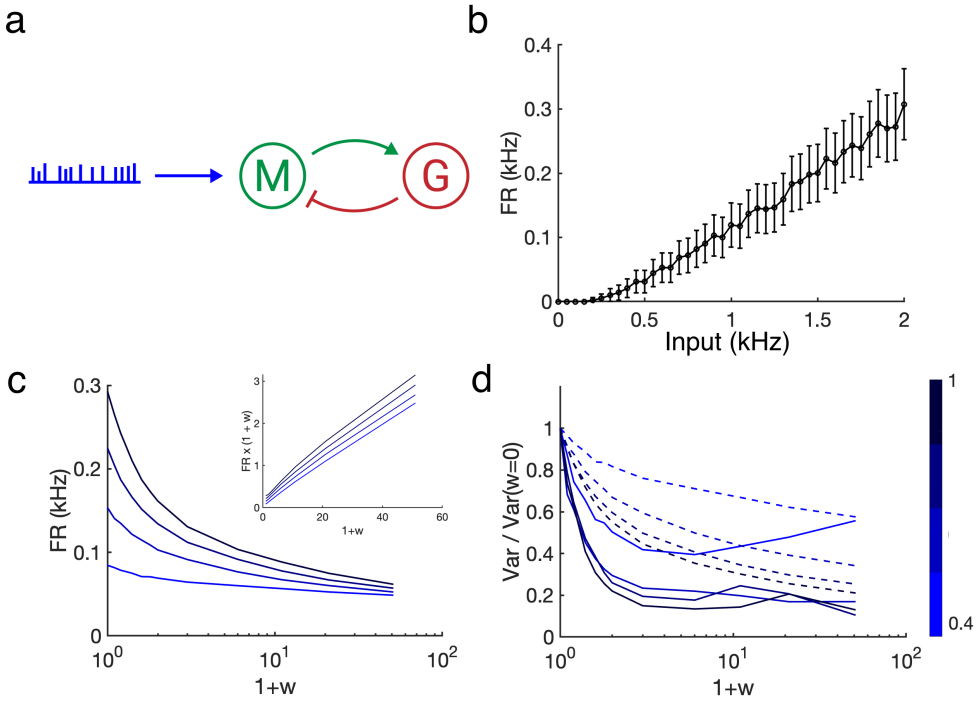


Figure 2.1. Properties of a reciprocally-connected MC-GC pair. (a) Network schematic. (b) FI curve of exponential integrate and fire neuron. Horizontal axis is the rate of a Poisson spike train who provides input to the neuron, and vertical axis is firing rate. Both firing rates are measured in $kHz$. Error bars indicate standard deviation of spikes measured in $dt = 0.02$ iterations over a 2000 ms simulation. (c) Average firing rate with inhibitory strength $w$. Increasingly darker shades of blue correspond to increasing $w$. Inset of (c) shows FR multiplied by $(1+w)$. (d) Variance of rates (proportional to that of $w = 0$). Dashed lines show corresponding variances of Poisson spike trains of the same mean rates. Variance of rates is measured at $T = 60$.

spiking pattern (Figure 2.3ac), a rhythm emerges at sufficiently strong inhibitory coupling and strong inputs. These rhythms comprise alternating cycles of MC spiking that is silenced by GC spiking (Figure 2.3b), and the corresponding power spectral density for these MCs show strong peaks at the dominant system frequency and all its harmonics (Figure 2.3c). Small $T$ measurements (that is, smaller than the system period) pick up either high amounts of activity within a spiking volley or periods of silence based

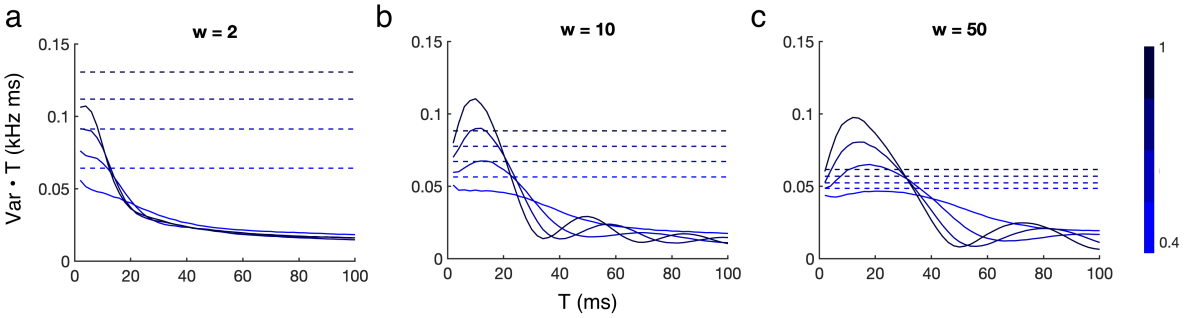


Figure 2.2. Variance of rates scaled by measurement duration for (a) $w = 2$, (b) $w = 10$, (c) $w = 50$. Dashed lines correspond to scaled variance of rates for equivalent Poisson trains. Increasingly darker shades of blue indicate increasing input strength.

off chance, resulting in high variability across measurements. Conversely, measurement windows near the period size pick up more consistent numbers of spikes, resulting in a lower estimated variance.

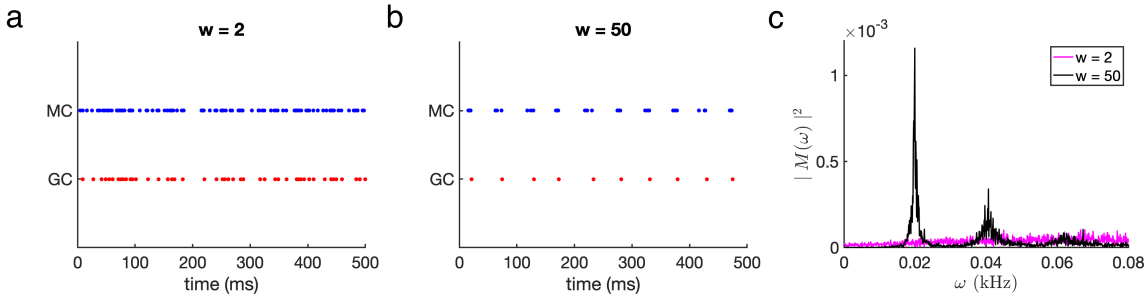


Figure 2.3. (a) Raster plot of MC (blue) and GC (red) for $w = 2$ during a randomly chosen 500 ms interval. (b) Raster plot for $w = 50$. (c) Power spectral density for $w = 2$ (magenta) and $w = 50$ (black). Relative input strength is 1.

## 2.2. Shared inhibition

We next slightly expand our model to assess the impacts of shared connections between multiple MCs. These simulations arose from the initial puzzlement that the 2-MC network behaved differently than 10-MC networks with the same parameters (Appendix A.4).

We simulate reciprocal networks containing a single inhibitory GC and increasing numbers of MCs. Each MC is excited by an independent Poisson input with fixed mean rate. Each MC excites the single GC and is inhibited by it. We use inhibitory strength $w = 50$ and take randomly placed measurement windows of size $T = 200$. A key distinction from the previous scenario is that these neurons are not independent; whereas a single mitral cell reciprocally connected to a single granule cell receives inhibition resulting from only its own activity, these mitral cells receive inhibition resulting from both its own activity and the collective activity of all its neighbors.

Indeed, the GC's variance reduction is lessened with increasing number of interconnected MCs (Figure 2.4a). Additionally, as the system size increases, any pair of mitral cells has a smaller pairwise correlation with each other (Figure 2.4b), as well as a smaller pairwise correlation with the granule cell (Figure 2.4c). Increasing the number of cells in an interconnected network dilutes the noise-shaping potency of the inhibitory signal.

Note that although variance does not decrease with $w$ at all window sizes (seen in Figure 2.3), increasing number of MCs increases the variance at any fixed $w$ and $T$.

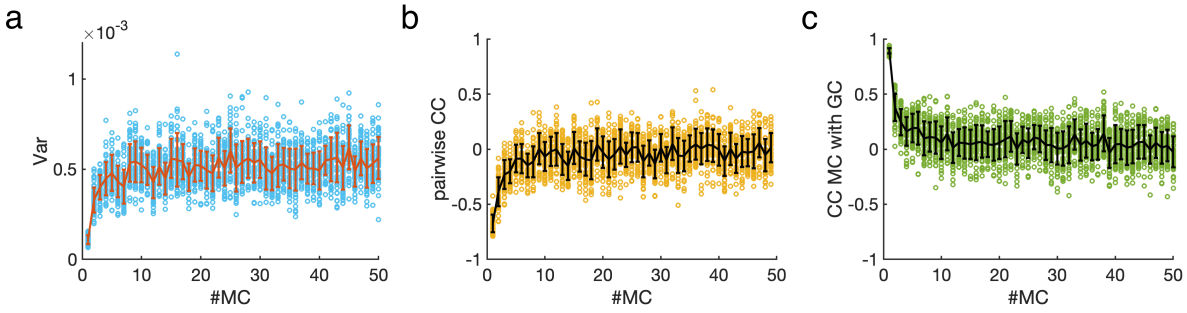


Figure 2.4. (a) Variance of MC rates, not scaled with $T$. (b) Pairwise correlation coefficient between MCs. (c) Pairwise correlation coefficient of MCs with the GC. Dots are data from individual samples. Solid line and error bars indicate mean and standard deviations, respectively.

## 2.3. Pairwise covariance

### 2.3.1. Pairwise covariances in a globally-connected network adhere to emergent rhythms.

To further explore how covariance terms are shaped by shared inhibition and measurement duration, we simulate all-to-all connected networks with at least 2 MC. In these networks, the number of GC is always equal to the number of MC, and every MC is reciprocally connected to every GC. To maintain each cell's input comparable to that of the single MC-GC network, we divide both coupling matrices $W_{mg}$ and $W_{gm}$ by the number of recipient cells. Note that since every granule cell receives input from all mitral cells, and there is no noise input to the granule cells, these systems are functionally equivalent to systems containing only one universally-connected granule cell. System equations are similar to that of the single MC-GC network:

$$\tau_{\mathrm{MC}_k}\frac{d}{dt}\bar{x}_{\mathrm{MC}_k} = -\bar{x}_{\mathrm{MC}_k} + \sum_{\mathrm{MC}_k \text{ spike}} \delta(t - t_{\mathrm{spike}}) \tag{2.5}$$

$$\tau_{\mathrm{GC}_i}\frac{d}{dt}\bar{x}_{\mathrm{GC}_i} = -\bar{x}_{\mathrm{GC}\ i} + \sum_{\mathrm{GC}_i \text{ spike}} \delta(t - t_{\mathrm{spike}}) \tag{2.6}$$

$$I_{\text{to MC}} = -\left(\sum_i w \times W_{gm} \times \bar{x}_{\mathrm{GC}\ i}\right) + \left(\sum_j W_{om} \times \bar{x}_{\mathrm{OSN}\ j}\right) \tag{2.7}$$

$$I_{\text{to GC}} = \left(\sum_k W_{mg} \times \bar{x}_{\mathrm{MC}\ k}\right). \tag{2.8}$$

Motivated by our previous finding that larger stimulus intensity drives stronger rhythms (Figure 2.2), we now stimulate an all-to-all network in response to two input strengths, with half the MCs driven by 1 kHz spike trains and the other half driven by 0.77 kHz spike trains. These values are arbitrarily chosen, and using two levels of stimulation sets

the stage for a later section in this thesis, in which we define a two-class discrimination problem between stimuli each defined by two levels of input.

The precise scaling of variance with input strength is difficult to quantify and varies with $T$ and $w$ (a few examples for the single-reciprocal scenario are shown in Figure 2.2), but we generally find that MCs receiving larger stimulation (Figure 2.5, top row, green curves) have larger variance than MCs receiving lower stimulation (Figure 2.5, top row, red curves). Variance also increases with larger system sizes (Figure 2.5).

At small $w$, both scaled variance and covariance monotonically decrease with $T$. Covariance in particular is negative, corresponding to MCs' indirectly suppressing each other through mutual shared inhibition. At larger $w$, variance and covariance exhibit large local maxima at small windows, indicating the presence of a system rhythm. Covariance crosses zero at $T$ near the system period, and the zero-crossing moves towards larger $T$ at stronger $w$ (Figure 2.5). System rhythms become visually apparent around $w = 2$, and synchronous volleys grow further apart as the inhibitory coupling strength increases (Figure 2.6). The power spectral density exhibits peaks that increase in amplitude as $w$ increases, indicating stronger rhythms with larger inhibition.

Visual features of covariance are consistent with experimental and computational results [**23**]. Positive covariance at small $T$ indicates correlated firing, such as in synchronous volleys of spikes. Negative covariance at larger $T$ is consistent with MCs' suppressing each other indirectly through shared inhibition. Moreover, covariance slightly increases towards zero for larger system sizes, suggesting that negative feedback is less potent when diluted by the activities of more cells.

Note: At $w = 0$, variance is mostly constant across window size, with a small uptick at small window sizes. Since there is no interaction between mitral and granule cells in this scenario and the uptick does not decrease with more samples, this may be due to increased variability in the system on small time scales, for instance caused by the relative refractory period caused by resetting voltage after each spike.

Note: In this case, this uptick does not seem to be due to measurement error, but there are other scenarios in which measurement error scales inversely with window size. In Appendix B.1, we derive how linear discriminability in Poisson networks is overestimated by an amount that scales linearly with the measurement duration.

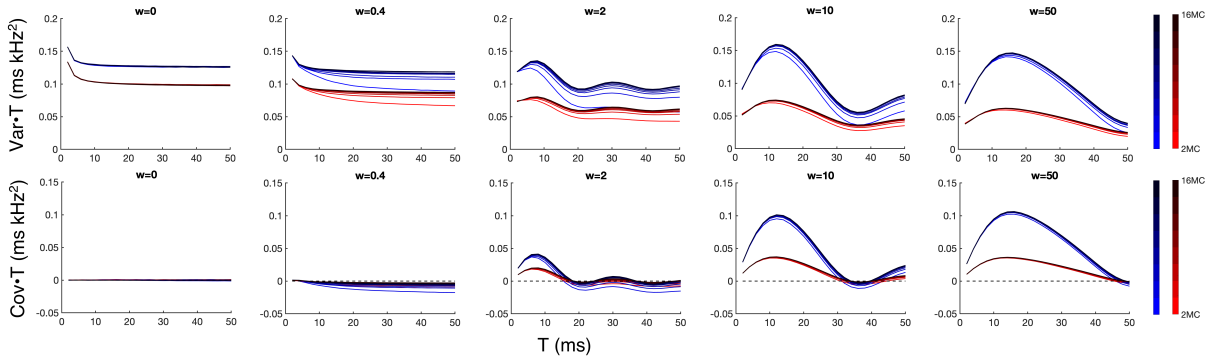


Figure 2.5. Variances and covariances of two cells with shared inhibition, stimulated by 1 kHz spike trains (blue) and 0.77 kHz spike trains (red). (Top row) Average variance scaled by $T$. (Bottom row) Pairwise covariance scaled with $T$. Increasingly darker shades indicate increasing system sizes. Dashed black line indicates zero.

### 2.3.2. Input strength determines magnitudes of pairwise covariance.

In subsequent sections of stimulus discrimination, we consider stimulus pairs that are symmetric under the exchange of MCs: a MC that receives $X$ Hz from Stimulus 1 and $Y$ Hz from Stimulus 2 has a corresponding MC in the network that receives $Y$ Hz from

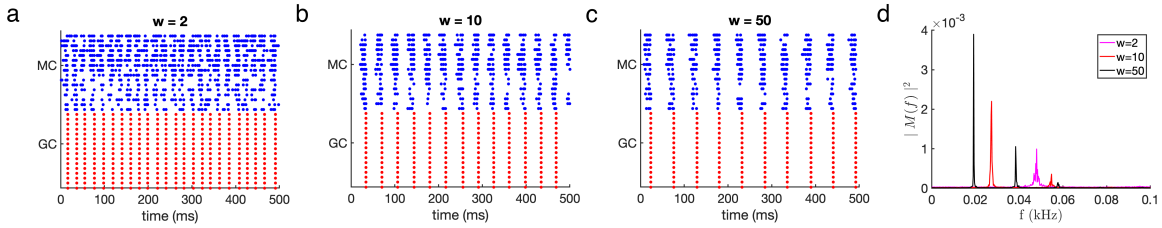


Figure 2.6. Raster plots for an all-to-all connected network with half the MCs driven by 1 kHz spike trains (blue) and half driven by 0.77 kHz spike trains (red), for (a) $w = 2$, (b) $w = 10$, (c) $w = 50$ during a randomly selected 500 ms interval. (d) Power spectral density for 16 MCs averaged over all mitral cells in the system.

Stimulus 1 and $X$ Hz from stimulus 2. These pairs may represent binary mixtures, which are often used in experimental studies [**27, 28, 29**]. Here, we briefly mention how the pairwise covariance of two cells is shaped by stimulus profiles. Consider the following two pairs of stimuli, termed High-Contrast and low-contrast blocks (Figure 2.7).

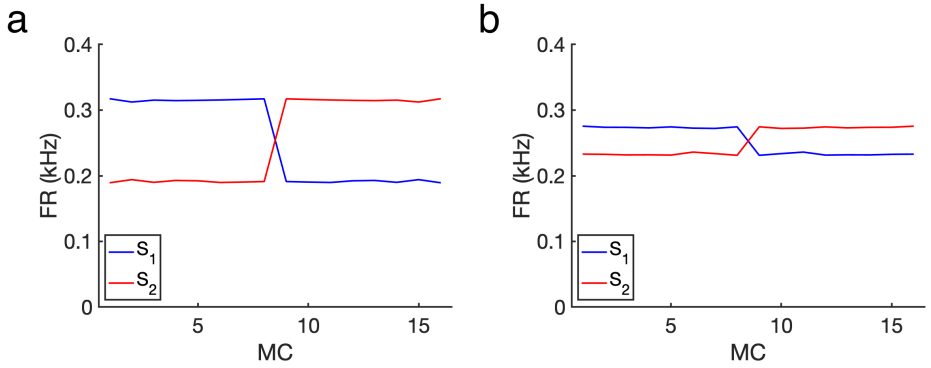


Figure 2.7. Two examples of symmetric block stimuli with the same mean firing rates with (a) high contrast and (b) low contrast. Data shows average firing rate over a 50,000 ms simulation.

Any two identically-stimulated MCs covary more strongly with larger stimulus intensity (Figure 2.8 top and bottom rows, colored dashed lines). However, the covariance of two differently-stimulated cells is not equal to the average covariance of two high-stimulated cells and two low-stimulated cells. Rather, MCs receiving different inputs in

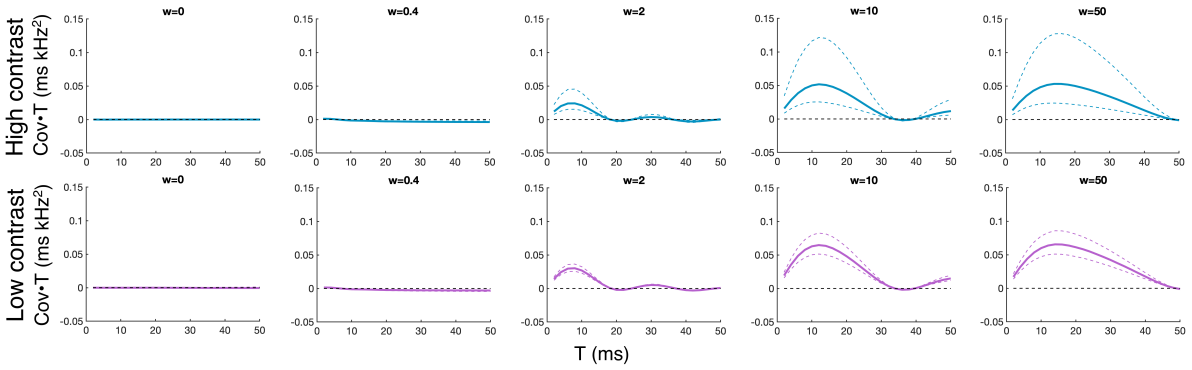


Figure 2.8. Scaled pairwise covariances of MCs in response to high-contrast blocks (top row) and low-contrast blocks (bottom row). Higher dashed line corresponds to two cells receiving the higher input; lower dashed line corresponds to two cells both receiving the lower input. Solid line corresponds to two cells receiving different inputs. Dashed black line is zero. Pairwise covariances are averaged over all pairs of cells in each category.

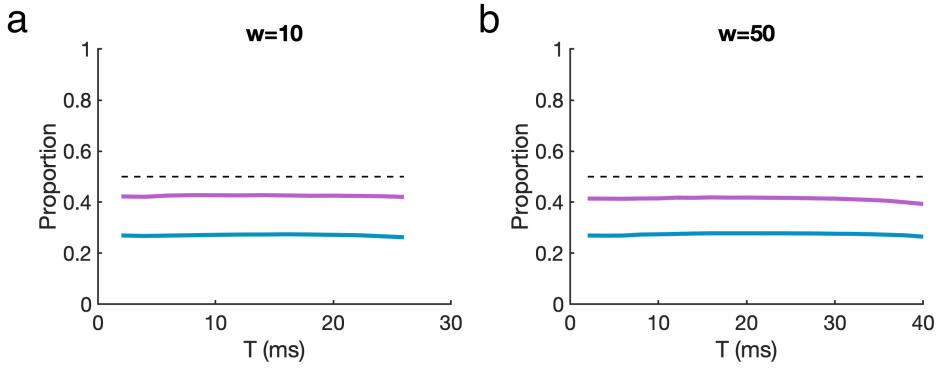


Figure 2.9. $(Cov_{high,low} - Cov_{low})/(Cov_{high} - Cov_{low})$ for $w = 10$ and $w = 50$. Purple line corresponds to 'low-contrast blocks and teal line corresponds to high-contrast blocks.

low-contrast blocks have a higher covariance than those of corresponding MCs in high-contrast blocks (Figure 2.8 top and bottom rows, solid colored lines) and covariance is closer to the average of its corresponding identically-stimulated MCs (Figure 2.9). See Appendix A.6 for a full explanation.

CHAPTER 3

# Network connectivity impacts linear discriminability performance.

Linear discriminability is a useful metric for stimulus dissimilarity that has been used extensively in the literature [**17, 30, 14, 21**]. It is well-suited for the olfactory system, in which downstream piriform cortical neurons integrate collective MC activity. This functionality suggests an implicit "weighting" of olfactory bulb inputs, which we represent computationally as assigning numerical weights to each dimension of an activity vector.

$\mathcal{F}(\vec{c})$ is the response of a linear readout with weights given in $\vec{c}$.

$$\mathcal{F}(\vec{c}) = \frac{\left(\vec{c}\cdot\left(\langle\vec{M}^{(1)}\rangle - \langle\vec{M}^{(2)}\rangle\right)\right)^2}{\vec{c}^T\left(\sum^{(1)} + \sum^{(2)}\right)\vec{c}} \tag{3.1}$$

In this formulation, $\langle\vec{M}^{(i)}\rangle$ is a length-$N$ vector containing the average responses (in spikes) to stimulus $i$, and $\Sigma^{(i)}$ is the corresponding covariance matrix. Note that the magnitude of $\vec{c}$ is divided out; $\mathcal{F}(\vec{c})$ only depends on its direction.

Note also that for any $\vec{c}$, $\mathcal{F}$ increases with the magnitudes of responses. For a Poisson network for instance, in which covariance of spikes $\Sigma^{(i)}$ is a diagonal matrix whose entries correspond to mean spike counts, $\mathcal{F}$ increases linearly with measurement duration. This implies that a network facing a two-class discrimination problem can always improve its performance by measuring for longer. To account for this known improvement, our results report $\mathcal{F}$ divided by $T$, which represents a scaled metric.

For notational convenience, we denote $\Delta\vec{\mu}$ as the difference in average responses and $\Sigma$ as the sum of covariances, respectively:

$$\Delta\vec{\mu} = \langle\vec{M}^{(1)}\rangle - \langle\vec{M}^{(2)}\rangle \tag{3.2}$$

$$\Sigma = \Sigma^{(1)} + \Sigma^{(2)}, \tag{3.3}$$

The optimal vector $\vec{c}_{opt}$ can be found by taking the Fréchet derivative of $\mathcal{F}(\vec{c})$, setting equal to zero, and solving for $\vec{c}$.

$$\vec{c}_{opt} \propto \Sigma^{-1}\Delta\vec{\mu}, \tag{3.4}$$

and as a result,

$$\mathcal{F}_{opt} = \Delta\vec{\mu}^T\Sigma^{-1}\Delta\vec{\mu}. \tag{3.5}$$

We also note that $\mathcal{F}_{opt}$ can be written in terms of the eigendecomposition of $\sum$,

$$\mathcal{F}_{opt} = \sum_j \frac{1}{\lambda_j}\left(\Delta\vec{\mu}\cdot\vec{e}_j\right)^2, \tag{3.6}$$

where $\lambda_j$ and $\vec{e}_j$ are the $j$th eigenvalue and normalized eigenvector of $\sum$, respectively.

### 3.1. Symmetric block stimuli

Consider a network exposed to symmetric block stimuli: half the MCs in the system receive input from independent Poisson spike trains of one mean rate, and the other half of the MCs receive a constant input from independent Poisson spike trains of the other mean rate. In the infinite-time limit, the average MC firing rates of the network

approximates a symmetric block shape (Figure 3.1) whose average difference $\Delta\mu$ is odd under the exchange of mitral cells and is governed by a single number: the absolute difference in response rates between MCs in different groups (Figure 3.1b).

If the system is single- or globally-connected, the expected sum of covariances $\Sigma$ contains three distinct terms (Figure 3.1c). It has the form

$$\Sigma = \begin{bmatrix} a & b & \dots & b & c & c & \dots & c \\ b & a & b & \vdots & c & c & c & c \\ \vdots & b & \ddots & b & \vdots & c & \ddots & \vdots \\ b & \dots & b & a & c & c & \dots & c \\ c & c & \dots & c & a & b & \dots & b \\ c & c & c & c & b & a & b & \vdots \\ \vdots & c & \ddots & \vdots & \vdots & b & \ddots & b \\ c & c & \dots & c & b & \dots & b & a \end{bmatrix}, \tag{3.7}$$

where $a$ is the sum of variances, $b$ is the sum of covariances of two identically-stimulated cells, and $c$ is the sum of pairwise covariances between two cells receiving different stimulation. The sum is taken over both stimuli.

For an $N \times N$ matrix, a straightforward calculation shows that out of $N$ mutually orthogonal eigenmodes, only one is parallel to $\Delta\vec{\mu}$, and it has eigenvalue $\left[a + \left(\frac{N}{2} - 1\right) b - \frac{N}{2} c\right]$.

$$\vec{e} = \underbrace{[1, 1, ..., 1, -1, -1, ..., -1]}_{N}{}^{T} \tag{3.8}$$

$$\sum \times \vec{e} = \left[a + \left(\frac{N}{2} - 1\right) b - \frac{N}{2} c\right] \vec{e}. \tag{3.9}$$

$\mathcal{F}_{opt}$ thus takes the form

$$\mathcal{F}_{opt} \propto \frac{1}{a + \left(\frac{N}{2} - 1\right) b - \frac{N}{2} c}. \tag{3.10}$$

Note that if $N = 2$, that

$$\mathcal{F}_{opt} \propto \frac{1}{a - c}. \tag{3.11}$$

In order to isolate the contributions of pairwise covariances in discriminability, we frequently compare $\mathcal{F}_{opt}$ with $\mathcal{F}_{uncorr}$, an equivalent calculation that takes $b = c = 0$.

$$\mathcal{F}_{uncorr} \propto \frac{1}{a}. \tag{3.12}$$

Figure 3.1 shows an example of numerically-estimated firing rates, $\Delta\vec{\mu}$, and covariance terms of an all-to-all connected network exposed to block stimuli.

## 3.2. 1:1 connectivity

We simulate a network of 16 MC each connected to its own GC (akin to Section 2.1). There is no lateral interaction, so pairwise covariances are zero, and discriminability is determined by average difference in response $\Delta\vec{\mu}$ and variance terms $a$. This connectivity is also referred to as "single connectivity."

In a series of simulations, we keep all parameters identical except for $w$, the strength of inhibition between each granule cell and mitral cell. As expected, overall firing rates of all 16 mitral cells decreases as the strength of inhibitory coupling decreases (Figure 3.2b). Furthermore, mitral cells who receive higher stimulation experience a greater decrease

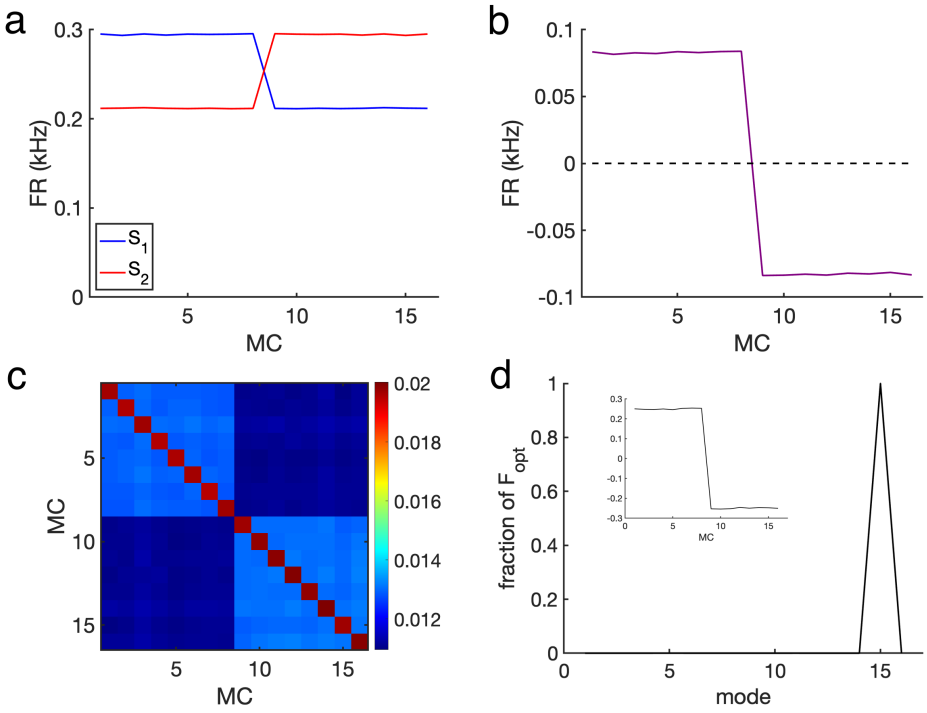


Figure 3.1. Example measurements of an all-to-all connected network exposed to symmetric block stimuli. (a) Average MC firing rates in response to two symmetric block stimuli (red and blue) at $w = 0$. (b) Difference in average firing rates $\Delta\vec{\mu}$. (c) Numerically-estimated covariance $\Sigma$ exhibits three distinct quantities, which we denote $a$, $b$, and $c$. $\Sigma$ is a sum over both stimuli. Data shown is from inhibition level $w = 50$ and measurement duration $T = 12$ ms. (d) Fraction of $\mathcal{F}_{opt}$ corresponding to each eigenmode of $\Sigma$, at $w = 50$. Inset shows eigenvector 15.

in activity, and the magnitude of $\Delta\vec{\mu}$ correspondingly decreases as inhibitory coupling increases (Figure 3.2cd).

To explicitly show how the scaled sum of variances changes with inhibition and window size, we plot traces of $a, b$, and $c$, as well as the sum $a+\left(\frac{N}{2}-1\right)b-\left(\frac{N}{2}\right)c$ across measurement durations $T$ from 2 to 50 milliseconds (Figure 3.3, top row).

The sum of variances $a$ trends downward with increased inhibitory coupling, but does not decrease monotonically with window size. In particular, for $w = 10$ and $w = 50$, $a$

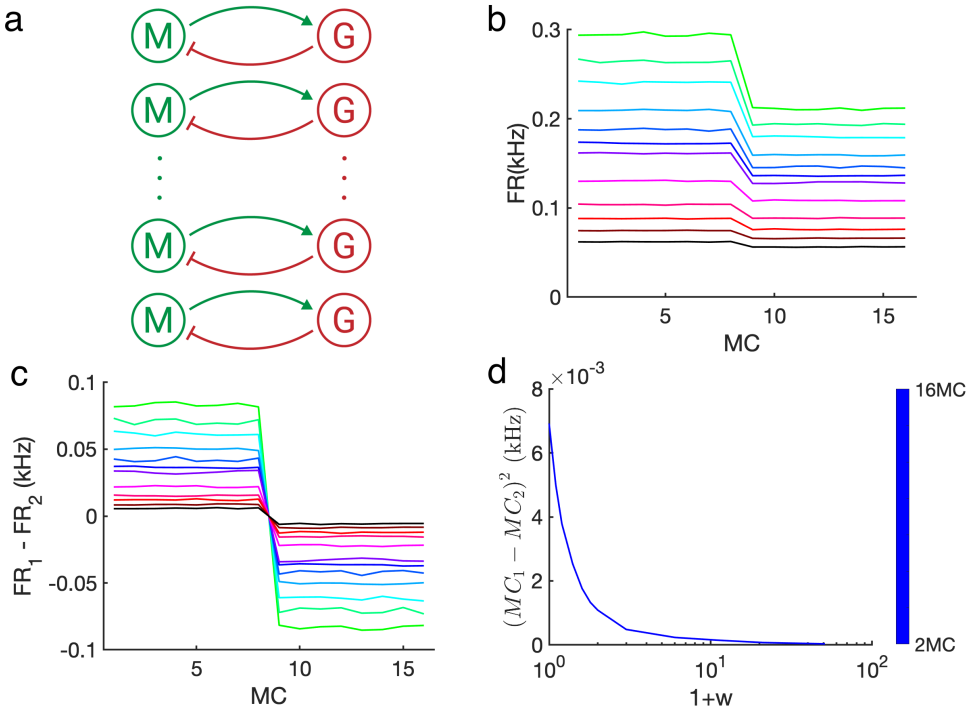


Figure 3.2. Network with 1:1 connectivity (also known as single connectivity) exposed to symmetric block stimuli. a) Schematic of network. Green arrows are excitatory synapses and red capped bars are inhibitory synapses. Not pictured: incoming excitatory synapses from olfactory sensory neurons and outgoing excitatory synapses to downstream piriform cortex neurons. b) Firing rates in response to stimulus 1. Colors indicate varying inhibitory strengths from $w = 0$ to $w = 50$. c) Average difference between stimuli 1 and 2. Colored bands indicate varying inhibitory strengths and correspond to the firing rate curves in 3*b*. d) Squared average magnitude of $\Delta\tilde{\mu}$, averaged across all mitral cells.

exhibits a rise at small window sizes before decreasing. Nonmonotonicity in $a$ becomes highlighted in the inverse terms $\frac{1}{a}$, which trend upwards as inhibition coupling increases (left to right columns) but exhibits large nonmonotonicities with $T$ (Figure 3.3, middle row). Covariance terms are zero (Figure 3.3 top row, green and blue lines).

$\mathcal{F}_{opt}$, trends downwards with inhibition and upwards with window size at each inhibition (Figure 3.3 bottom row). Since $\mathcal{F}_{opt}$ is proportional to the inverse sum of variances

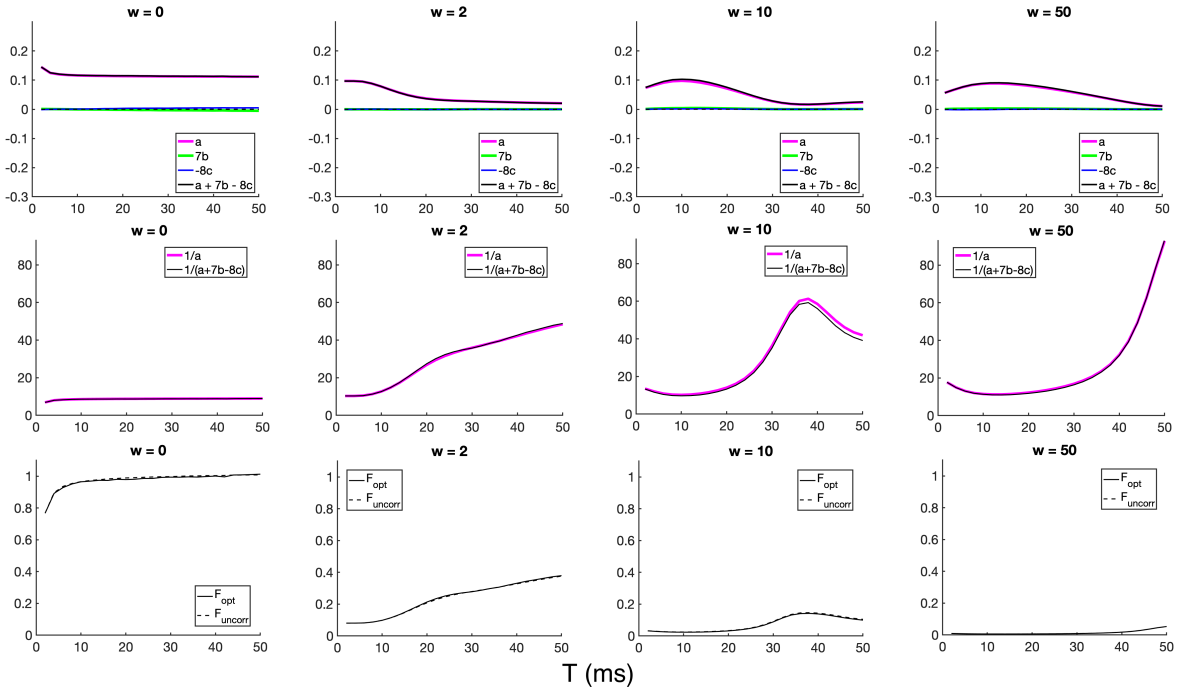


Figure 3.3. Variance contributions in a 1:1 connected network exposed to block stimuli. Top row: $a$ (magenta), $7b$ (green) and $-8c$ (blue). Solid black line is the sum of all three terms. Dashed black line indicates zero. Middle row: Inverses of $a$ (magenta) and $(a + 7b - 8c)$ (black). Bottom row: $\mathcal{F}_{opt}$ (solid black line) and $\mathcal{F}_{uncorr}$ (dashed black line), both scaled with $T$.

and proportional to the squared magnitude of $\Delta\vec{\mu}$, its downward trend must result because the reduction in $\Delta\vec{\mu}$ dominates over the marginal benefits of pairwise covariances. As expected, $\mathcal{F}_{opt}$ does not differ from $\mathcal{F}_{uncorr}$ because pairwise covariances are zero.

Notably, $\mathcal{F}_{opt}$ does not decrease with inhibition (Figure 3.4a) as quickly as the squared contribution of $\Delta\vec{\mu}$ (Figure 3.4d), suggesting that neurons' variances of rates are reduced faster than the squared contribution of their means. Furthermore, $\mathcal{F}_{opt}$ does not decrease with $w$ as steeply as its corresponding Poisson network at all window sizes, and its marginal improvement is larger at larger $T$.

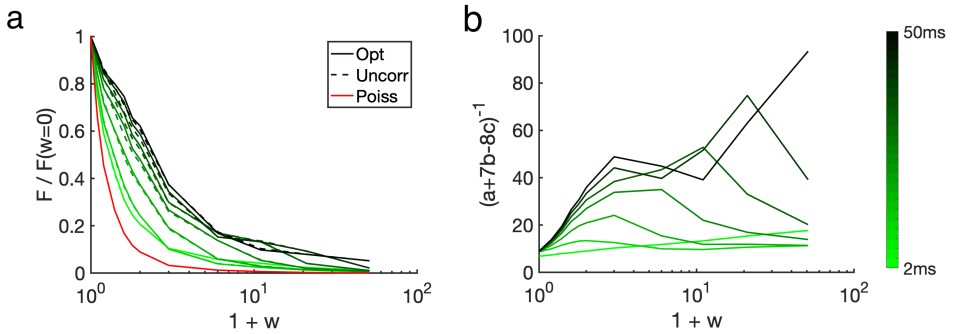


Figure 3.4. Optimal discriminability and variance contributions in a single-connected network. Increasingly darker of green correspond to increasing measurement durations from $T$ = 2 ms to 50 ms. (a) Optimal (solid lines) and uncorrelated (dashed lines) linear discriminability, scaled with $T$. Horizontal axis represents inhibitory coupling and is plotted on a log scale. We use $1 + w$ in order to show the zero weight on the graph. Red line indicates discrimination from a Poisson network with the same mean rates. (b) Contribution of terms from the covariance matrix, scaled by bin size $T$.

We conclude that inhibition in a single-connected network worsens discriminability, and the reduction is dominated by a reduction in the magnitude of $\Delta\vec{\mu}$.

### 3.3. Global connectivity

Next, we assess the linear discriminability of a network whose MCs and GCs are all-to-all (globally) connected (Figure 3.5a). In order to maintain each cell's input similar to that of the single connection case, we divide each connectivity matrix $W_{mg}$ and $W_{gm}$ by the number of their respective recipient cells. In Section 2.2, we found that as the number of MCs in a network increased, that MC variances were reduced less, that any given pair was less correlated with each other, and each MC was less correlated with the GCs. Motivated by this, we simulate globally connected networks containing a range from 2 MC to 16 MC.

In each of our networks of variable sizes, (Figure 3.5bd), MC input is set such that average firing rates are similar between systems of different sizes (e.g. MC 1 in a 2-MC network has the same average firing rate as MCs 1-8 in a 16-MC network). It follows that for each inhibition level, the corresponding element(s) of $\Delta\vec{\mu}$ is also identical (Figure 3.5ce). The squared magnitude of $\Delta\vec{\mu}$ decreases with inhibition just as in Single Connections, but less steeply at small inhibitions (Figure 3.6).

Like in our analyses of single-connected networks, we plot how the $\Sigma$ terms $a$ and $b$ vary with measurement duration $T$, at various inhibitions, and at various system sizes. Both $a$ and $b$ are qualitatively similar to before: monotonically decreasing with $T$ at small inhibitions, and with a local maximum emerging at larger $w$ (Figure 3.7, first two rows). Moreover, as system size increases, both variance and covariance trend upwards, indicative of larger system sizes' reduced ability to reduce noise.

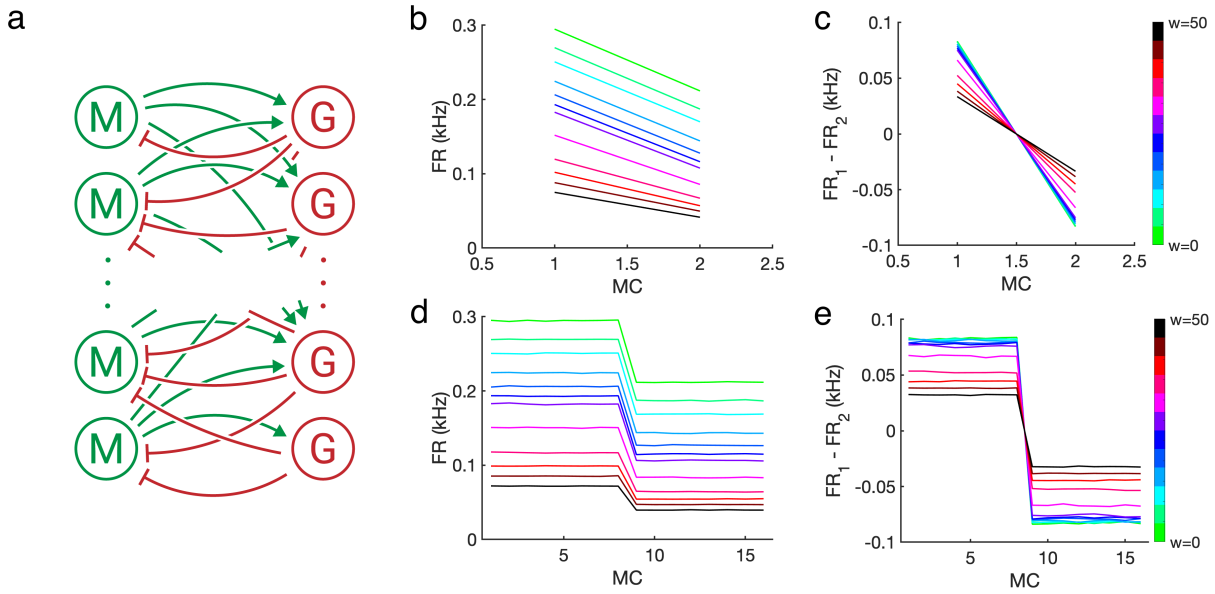


Figure 3.5. Firing rates and $|\,\Delta\vec{\mu}\,|^2$ in a globally connected network, $Nm = 2$ to $Nm = 16$. a) Schematic of Mitral-Granule coupling. In the simulated network, all mtiral cells excite all granule cells are also in turn inhibited by all granule cells. b) Average firing rates of two mitral cells and (c) corresponding signed differences in average firing rates. Different colored lines correspond from $w = 0$ to $w = 50$ (arbitrary units). d) and e) Same as b) and c) but for a network of 16 mitral cells receiving identical input.

The quantity $b - c$ is constant for all $N$. Even though $b$ and $c$ each increase with larger numbers of MCs, the difference of covariances between cells receiving the same stimulation and cells receiving different stimulation is consistent across different system sizes.

Recall that in a network exposed to symmetric block stimuli,

$$\mathcal{F}_{opt} \propto \frac{1}{a + \left(\frac{N}{2} - 1\right) b - \frac{N}{2} c}, \qquad \mathcal{F}_{uncorr} \propto \frac{1}{a}. \tag{3.13}$$

Smaller variance ($a$) benefits discriminability, and pairwise covariance terms $b$ and $c$ have opposing effects. Pairwise covariance term $b$ improves discriminability when it is negative, whereas pairwise covariance term $c$ improves discriminability when it is positive. In other

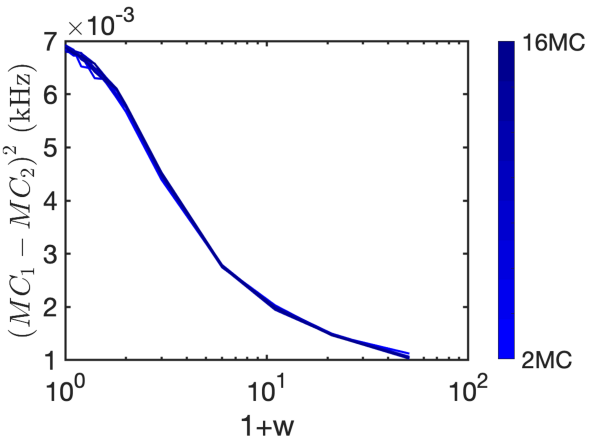


Figure 3.6. Squared average magnitude of $\Delta\vec{\mu}$, averaged across all mitral cells. Subscripts indicate stimulus index. Increasingly darker shades of blue correspond to larger system sizes.

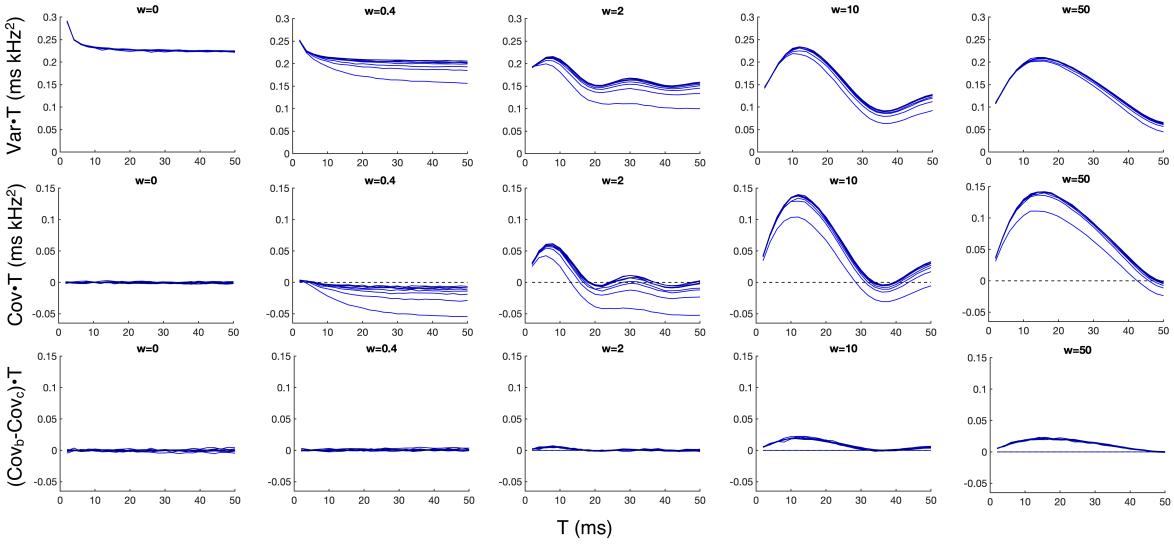


Figure 3.7. Scaled sum of variances and covariances in a globally connected network. (Top row) $a$ across measurement durations for different inhibitory coupling strengths. Increasingly darker shades of blue indicate increasing system sizes, as in Figure 3.6. (Middle row) Same as top row but for $b$. Dashed line indicates zero. (Bottom row) Same as first two rows but for the quantity $b - c$.

words, discriminability improves when cells receiving the same input are anti-correlated and cells receiving different inputs are positively correlated.

In Figure 3.8, we first plot linear discriminability terms $\mathcal{F}_{uncorr}$ and $\mathcal{F}_{opt}$, followed by the ratio $\mathcal{F}_{opt}/\mathcal{F}_{uncorr}$. The latter quantity divides out the contribution from $\Delta\vec{\mu}$ and indicates whether pairwise covariances improve discriminability.

$$\frac{\mathcal{F}_{opt}}{\mathcal{F}_{uncorr}} = \frac{a}{a + \left(\frac{N}{2} - 1\right) b - \frac{N}{2} c} \tag{3.14}$$

$$> 1 \text{ if pairwise covariances are beneficial.}$$

$\mathcal{F}_{uncorr}$ improves with small inhibition and more strongly in smaller system sizes (Figure 3.8ac). Since the contribution from $\Delta\vec{\mu}$ decreases monotonically with inhibition, this indicates that variance reduction is strong enough to compensate for the loss in average difference. In a small network, pairwise covariance benefits discriminability at smaller $T$ and sufficiently large inhibition. In a larger network (e.g. $N = 16$), pairwise covariance has a much smaller effect on discriminability.

In Section 2.3.2, we found that the amount of contrast between a stimulus pair alters the relative magnitudes of pairwise covariances. Recall that $b$ is the sum of covariances of identically-stimulated MC (sum of dashed lines in Figure 2.8), and $c$ is the sum of covariances of differently-stimulated MC (double the solid line of 2.8).

In a pair of stimuli of high contrast, $b$ is much larger than $c$, and

$$\left(\frac{N}{2} - 1\right) b > \frac{N}{2} c, \tag{3.15}$$

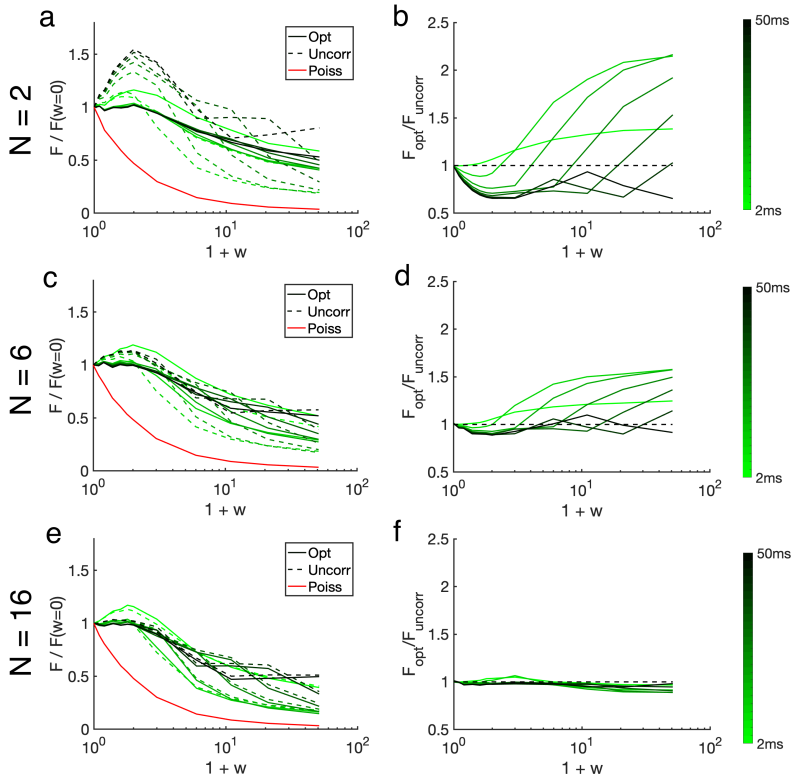


Figure 3.8. Discriminability in a globally-connected network, scaled with $T$. (a) 2 MC network. $\mathcal{F}_{opt}$ (solid lines) and $\mathcal{F}_{uncorr}$ (dashed lines) with inhibitory coupling $w$ (a.u.). Increasingly darker of green indicate larger measurement durations. Red line is the result of Poisson spike trains with the same mean rates. (b) Ratio $\frac{\mathcal{F}_{opt}}{\mathcal{F}_{uncorr}}$. (c) and (d) Same as (a) and (b) but for $N = 6$. (e) and (f) Same as (a) and (b) but for $N = 16$.

which reduces overall discriminability. Conversely, for stimuli of low contrast, $b$ is very close to $c$, and

$$\left(\frac{N}{2} - 1\right) b < \frac{N}{2} c, \tag{3.16}$$

which improves discriminability.

In short, pairwise covariance terms enhance discriminability when stimuli are similar to each other (low contrast), and they worsen discriminability when stimuli are far apart (high-contrast).

One functional implication of these data is that lateral inhibition induces correlations that are helpful for stimulus pairs that are very similar to each other, and that this pairwise covariance information harms discriminability for two stimuli with a large difference. There are studies suggesting that odor representations are reformatted differently depending on the difficulty of the discrimination task [**6**], and our work provides a concrete mathematical suggestion for these different modalities.

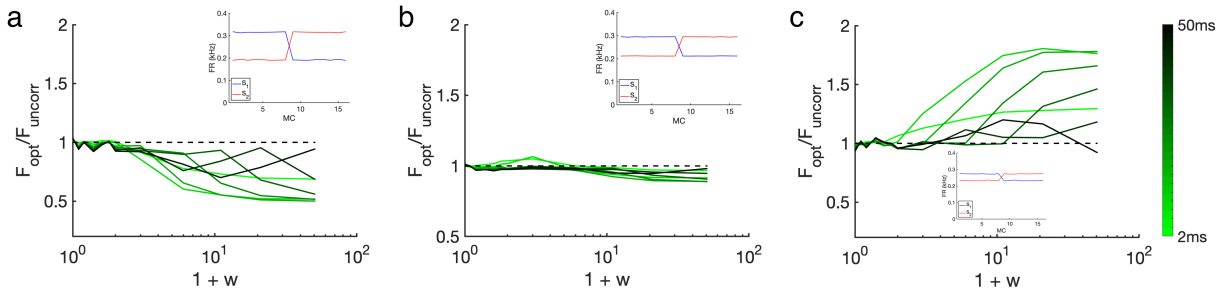


Figure 3.9. Proportion of $\mathcal{F}_{opt}$ to $\mathcal{F}_{uncorr}$ for different stimulus profiles, scaled with $T$. (a) high-contrast blocks (b) baseline stimulus contrast (c) low-contrast blocks. Increasingly darker shades of green indicate increasing measurement durations. Insets of (abc) indicate average MC firing rates.

### 3.4. Random (exploratory) linear discriminability

Up to this point, we have quantified network performance using optimal linear discriminability, which implicitly assumes that the network has presupposed knowledge how much to weight each MC input. Now we will consider a similar metric that better represents uninformed or exploratory performance, for instance when an animal is first exposed

to unfamiliar stimuli. We denote this metric $\mathcal{F}_{random}$ and compute it as follows.

$$\mathcal{F}_{random} = \frac{\left(\vec{c}_{rand} \cdot \left(\langle \vec{M}_1 \rangle - \langle \vec{M}_2 \rangle\right)\right)^2}{\vec{c}_{rand}^{\,T} \left(\sum_1 + \sum_2\right) \vec{c}_{rand}}, \qquad \vec{c}_{rand} = \left[x_1, ..., x_N\right]^T, \qquad x_i \sim iid\ N(0,1). \tag{3.17}$$

To estimate random discriminability, we take $10^6$ random samples of weights $\vec{c}_{rand}$, compute the resulting discriminability for each, and take the average across all samples. We found that using the median, rather than arithmetic mean, gives identical qualitative results but takes longer to compute. For that reason, we report the means only.

For single- and globally-connected networks exposed to symmetric block stimuli, $\langle \mathcal{F}_{random} \rangle$ behaves qualitatively similar to $\mathcal{F}_{opt}$. Specifically, in a single-connected network, $\langle \mathcal{F}_{random} \rangle$ trends downward for increasing inhibitory strengths (Figure 3.10ac). After dividing by $\mid \Delta\vec{\mu} \mid^2$ however, discriminability trends upwards with inhibition (Figure 3.10bd), indicating that much of the reduction in discriminability is due to the reduction of average difference, and that inhibition reduces the MC variance. In globally-connected networks, small inhibitions improve discriminability because noise reduction is able to overcome the smaller reduction in $\Delta\vec{\mu}$.

One key distinction of $\mathcal{F}_{random}$ is that it considers noise reduction in all directions, not only in the optimal direction (for more detailed visualizations, see Appendix A.9), and the two metrics differ when the stimuli drives a set of mitral cells at the same mean rates. In these stimuli (Figure 3.11), $\Delta\vec{\mu}$ is zero in the entries corresponding to identically-driven MCs. Single-connected networks exposed to such stimuli would have identical properties as if the middle distractor block were absent (because distractor MC have zero entries in $\Delta\vec{\mu}$, and noise correlations between cells in the wings are unaffected by the distractor set). In globally-connected networks exposed to such stimuli, activity of the middle distractor

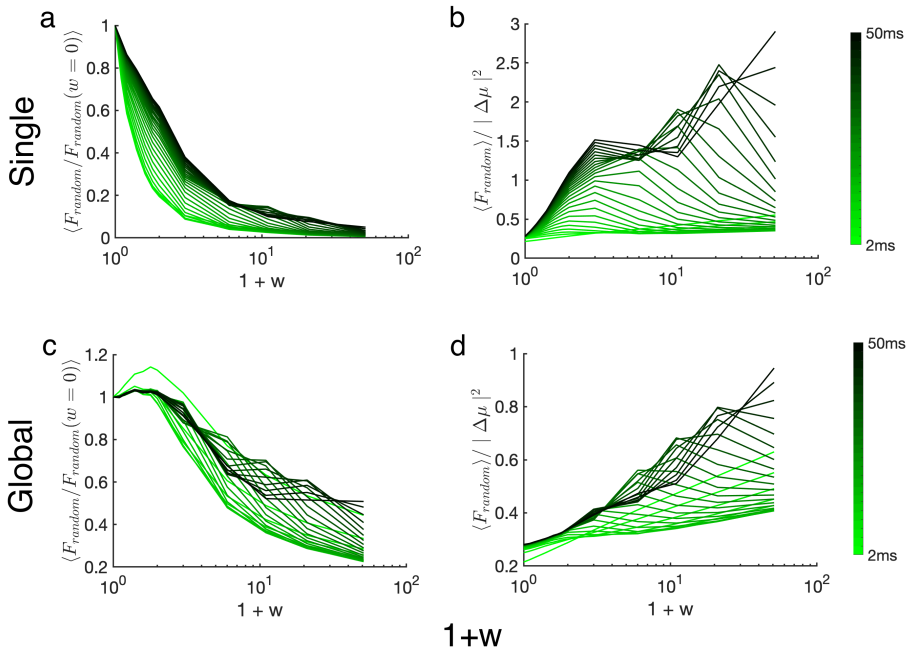


Figure 3.10. (a) Mean $\mathcal{F}_{random}$, scaled with $T$, over inhibitory strengths for single-connected network in response to block stimuli. Increasing shades of green corresponding to larger measurement durations. (b) Same as (a) but divided by the contribution from $\Delta\vec{\mu}$. (c) and (d) Same as (a) and (b) for a globally-connected network.

set could impact noise correlations of cells in the wings due to the shared inhibition, and we find that in stimuli with larger middle block stimulations, pairwise correlations within the wings worsen discriminability (Appendix A.7).

Inhibition that is neither 1:1 nor all-to-all, but rather selectively suppresses these distractor cells substantially improves $\mathcal{F}_{random}$. In these regimes, $\mathcal{F}_{opt}$ is unchanged by inhibiting this distractor set (optimal discriminability has no contribution from cells whose average difference is 0) but this selective reduction reduces the noise contributions from these distractor cells. Indeed, inhibition that selectively suppressing distractor cells (reducing their noise contributions) improves $\mathcal{F}_{random}$. Moreover, inhibition that selectively

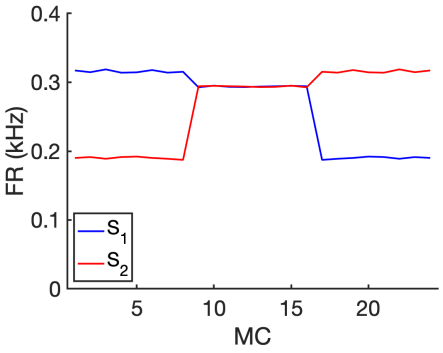


Figure 3.11. Firing rates of MCs exposed to symmetric block stimuli, with an additional 8 MC that fire at the same rate across both stimuli.

suppressing cells in the wings worsens $\mathcal{F}_{random}$ (Figure 3.12). This latter reduction is dominated by a decrease in $\Delta\vec{\mu}$; dividing out its contribution shows that noise is still reduced (Figure 3.13).

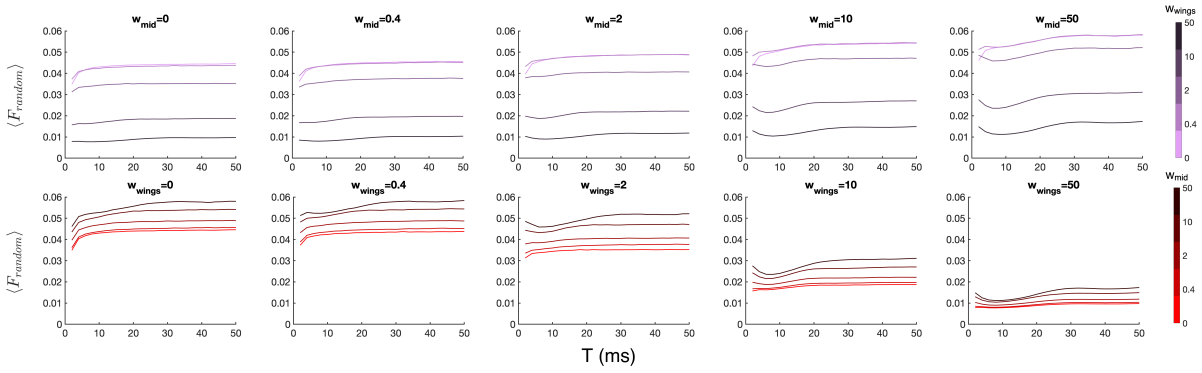


Figure 3.12. (Top row) Average $\mathcal{F}_{random}$, scaled with $T$, for a system exposed to high-contrast blocks with a middle distractor set. Left to right correspond to increasing levels of preferential distractor stimulation. Increasing shades of purple indicate increasing inhibition to cells in the wings. (Bottom row) Same as (top row), but left to right indicates increasing levels of inhibition delivered in the wings. Increasingly darker shades of red indicate increasing amounts of middle block inhibition.

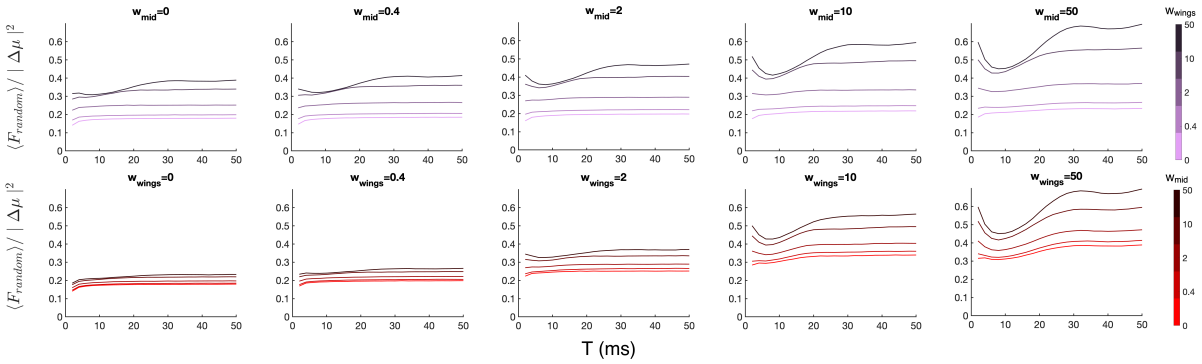


Figure 3.13. Same data as in Figure 3.12 but divided by the squared magnitude of $\Delta\vec{\mu}$

## 3.5. Corollary: generalizing to other stimuli

To this point, we have only considered stimuli that are block-shaped; we now demonstrate that the results from block stimuli extend to more general classes of odorants. The first of two new stimulus pairs is composed of two symmetric mixtures of Gaussian curves, and it represents pairs whose difference is largest in the MCs who are most strongly driven by each stimulus (Figure 3.14a). The second pair, composed of overlapping skewed Gaussian curves, represents classes of odorant pairs whose difference is largest in more weakly stimulated MCs. We refer to these stimuli as "Gaussian mixtures" and "skewed Gaussians," respectively. While these two pairs by all means do not represent all possible classes of odorant pairs, they serve as examples of inputs whose possible MC responses exist on a continuum.

For notational convenience, we borrow terminology from [**31**] and denote a stimulus' "primacy set" as the set of neurons that receive the most stimulation. Gaussian mixtures have different primacy sets; skewed Gaussians have the same primacy set.

### 3.5.1. Single and global connectivity

As in the block stimuli, inhibition delivered through single reciprocal connections decreases $\mathcal{F}_{opt} = \mathcal{F}_{uncorr}$ monotonically, but not as steeply as that of a corresponding Poisson network. Increasing window size also increases $\mathcal{F}_{opt}$ (scaled with $T$) (Figure 3.15).

In a network of all-to-all connections, small inhibition marginally improves discriminability in with small $T$, and pairwise covariances are more beneficial if the two Gaussian curves are closer together. (Figure 3.16abc).

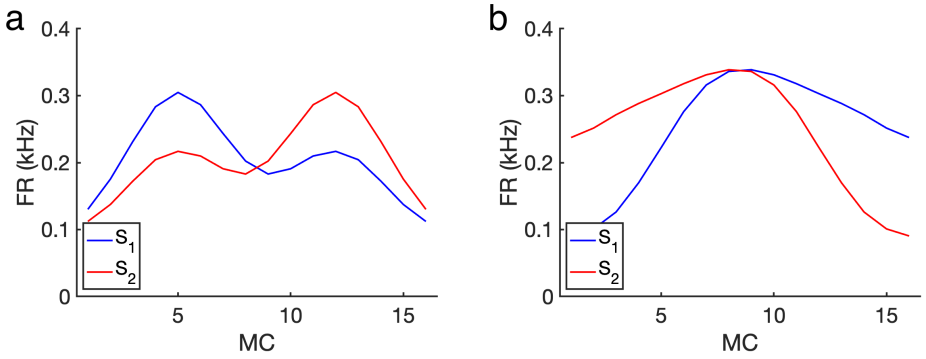


Figure 3.14. (a) Gaussian mixtures and (b) Skewed Gaussians. Vertical axis is average MC response.

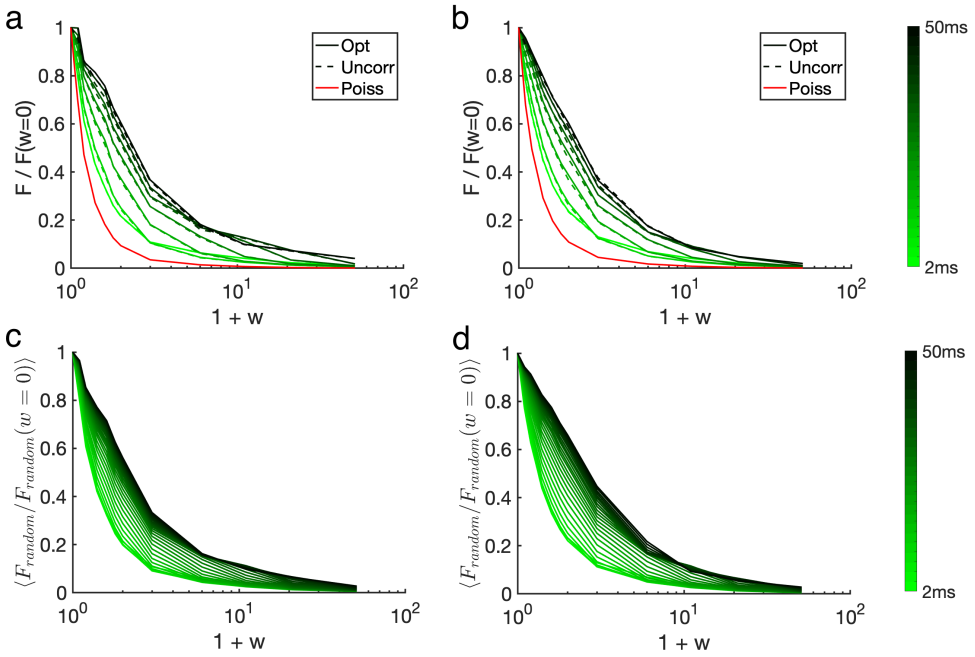


Figure 3.15. Top row: $\mathcal{F}_{opt}$ with inhibition in a single-connected network, relative to $w = 0$ and scaled with $T$, for (a) Gaussian mixtures and (b) Skewed Gaussians. Bottom row: Same as top row but for $\langle \mathcal{F}_{random} \rangle$

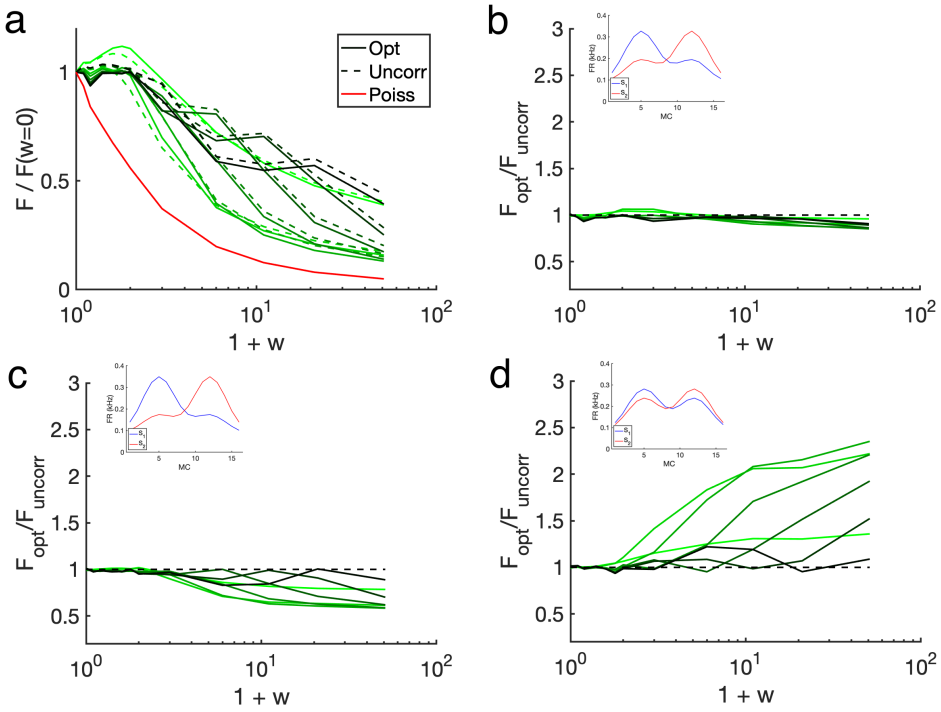


Figure 3.16. (a) $\mathcal{F}_{opt}$ with inhibition in a globally-connected network, relative to $w = 0$ and scaled with $T$, for Gaussian mixtures. (bcd) Proportion $\mathcal{F}_{opt}/\mathcal{F}_{uncorr}$ for various levels of contrast between stimuli. Insets show average MC response.

### 3.5.2. Stimulus-based inhibition

In section 3.4, we showed how exploratory discriminability can be improved by preferentially inhibiting MCs who did not spike differently across stimuli, reducing noise contributions from those cells without changing the average difference in response. The idea of these "preferential networks" that only inhibit a subset of the MCs is motivated by studies in the literature in which synaptic or structural plasticity occur as a result of some experience-driven learning. In these studies, connectivity may be shaped by previous exposure to sensory input, and the resulting changes in the network can improve

discriminability [**32, 33, 6, 34**]. We now similarly assess whether inhibition delivered through stimulus-dependent connectivity can improve discriminability.

Olfactory bulb physiology is characterized by long-ranging MC connections and short-ranged GC-MC reciprocal interactions, and axonless GCs deliver inhibition primarily through dendrodendritic synapses near MCs' soma [**5**]. To model the effects of "learned" connectivity, we leave the granule-to-mitral inhibitory matrix $W_{gm}$ as a diagonal matrix with identical elements and only modify the mitral-to-granule coupling matrix $W_{mg}$. Each entry in $W_{mg}$ is defined such that in the mitral-to-mitral coupling matrix $W_{mm} = W_{mg} \times W_{gm}$, the $i,j$th entry $W_{mm}^{i,j}$ is proportional to the activities of MC $i$ and MC $j$.

A key feature of stimulus-based connectivity is that MCs who are most highly stimulated are suppressed most strongly. For Gaussian mixtures, the mitral-to-mitral interaction matrix is heavily weighted on the MCs in the peaks (Figure 3.17b), resulting in firing rate profiles that are preferentially reduced in the peaks and a $\Delta\vec{\mu}$ that decreases in magnitude in those corresponding cells (Figure 3.17cd). Intuitively, this may substantially worsen discriminability of Gaussian mixtures, since $\Delta\vec{\mu}$ dominates linear discriminability, and we indeed find that $\mathcal{F}_{opt}$ and $\mathcal{F}_{random}$ both trend downwards with larger $w$ (Figure 3.19ac).

However, neither metric decreases as strongly as what might be expected from the reduction in $\Delta\vec{\mu}$ (Figure 3.19a, red line), suggesting that noise reduction offsets the overall worsening of discriminability. This is most noticeable at small inhibitions, for which $\Delta\vec{\mu}$ does not noticeably change (Figure 3.17d), and improvements in discriminability must be due to effects of the terms in $\sum$. While we cannot analytically derive the exact contributions these $\sum$ terms, we can conclude the following. First, inhibition reduces

variance, which enhances discriminability ($\mathcal{F}_{uncorr} > \mathcal{F}_{Poiss}$). Second, increasing inhibition induces pairwise covariances that are additionally beneficial (Figure 3.19b).

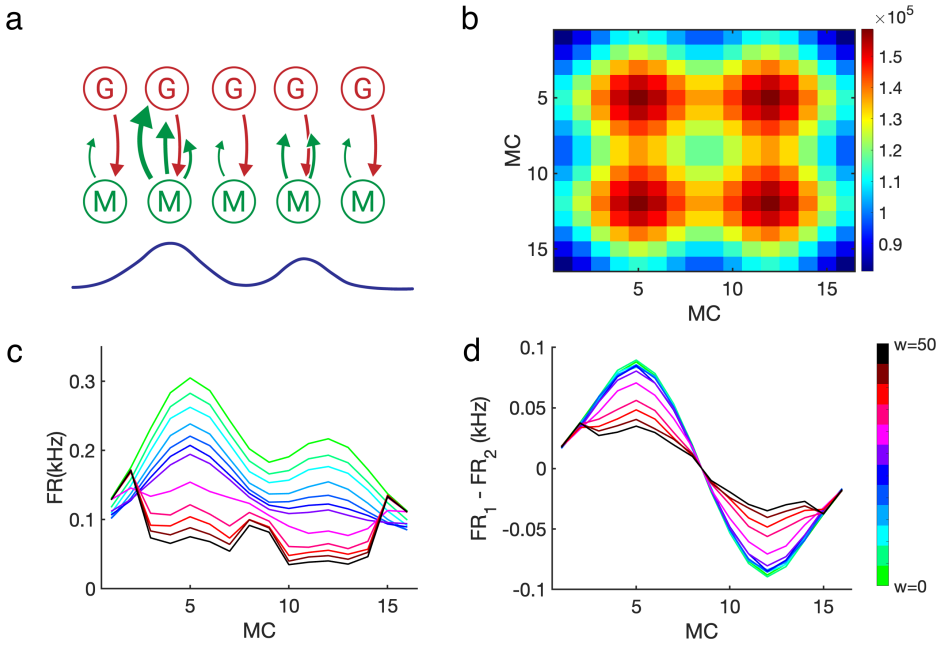


Figure 3.17. Stimulus-based recurrent inhibition and Gaussian mixtures. (a) Network schematic. (b) Product of MC-to-GC coupling matrix and GC-to-MC coupling matrix, representing the relative amount of inhibition the MCs indirectly deliver unto themselves. (c) Average firing rate in response to stimulus 1 for different levels of inhibition (green to black). (d) Difference in average firing rates $\Delta\vec{\mu}$.

A network exposed to skewed Gaussians contains a mitral-to-mitral interaction matrix that preferentially weights MCs in the center blob (Figure 3.18b) and preferentially silences the highly active cells in the middle (Figure 3.18). This is similar in idea to the setup in section 3.4, in which preferentially silencing distractor cells improved $\mathcal{F}_{random}$ by reducing noise from cells who fired similarly in both stimuli. Unlike in those previous analyses, stimulus-based inhibition of skewed Gaussians changes $\mathcal{F}_{opt}$, an effect that was absent when the stimuli were block-shaped. This is due to the continuous nature of the

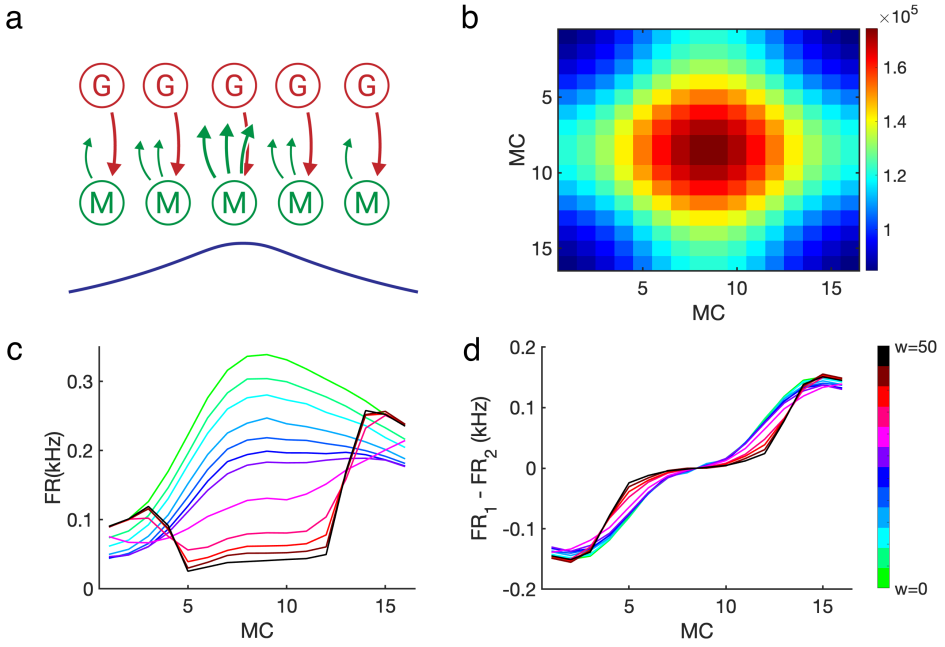


Figure 3.18. Stimulus-based recurrent inhibition and skewed Gaussians. (a) Network schematic. (b) Product of MC-to-GC coupling matrix and GC-to-MC coupling matrix, representing the relative amount of inhibition the MCs indirectly deliver unto themselves. ($W_{gm}$ is a diagonal matrix with identical entries) (c) Average firing rate in response to stimulus 1 for different levels of inhibition (green to black). (d) Difference in average firing rates $\Delta\vec{\mu}$.

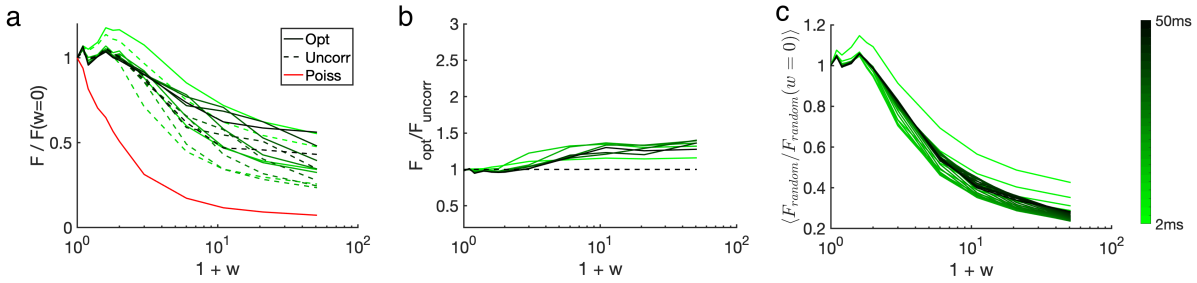


Figure 3.19. Discriminability scaled with $T$ of a stimulus-based connected network responding to Gaussian mixtures. (a) $\mathcal{F}_{opt}$ (b) $\mathcal{F}_{opt}/\mathcal{F}_{uncorr}$ (c) $\mathcal{F}_{random}$. Data in (ac) are taken with respect to the measurement at $w = 0$

stimuli (and therefore the coupling); suppression of the most active cells also suppresses neighboring cells whose differences contribute to $\mathcal{F}_{opt}$.

Notably, stimulus-based inhibition of skewed Gaussian stimuli significantly improves $\mathcal{F}_{random}$ (Figure 3.20c), and more strongly at small $T$. This suggests that a system learning to distinguish between two similar stimuli benefits from suppressing MCs that spike early and strongly, which do not contribute to stimulus difference but do contribute to total noise.

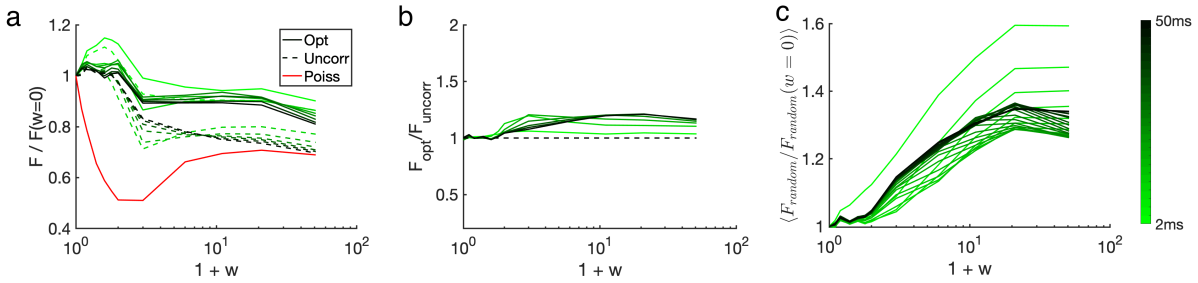


Figure 3.20. Discriminability scaled with $T$ of a stimulus-based connected network responding to skewed Gaussians. (a) $\mathcal{F}_{opt}$ (b) $\mathcal{F}_{opt}/\mathcal{F}_{uncorr}$ (c) $\mathcal{F}_{random}$. Data in (ac) are taken with respect to the measurement at $w = 0$

CHAPTER 4

# Timing-based metrics and the sniff cycle

So far, our model structure assumes that stimulus is continually present and responding cells have no time-based adaptation nor saturation. However, measurement duration can yield vastly different estimates of both average rate and variance (and consequently, discriminability) based on their size relative to the period of the system gamma rhythm.

Moreover, experimental evidence suggests that cell responses are locked to sniff onset and that precise temporal spiking patterns may inform towards odor discrimination [**35, 36, 31**], and recurrent connectivity in the cortex limits the amount of incoming information to just those coming from the earliest-spiking cells [**37**]. These studies suggest that discriminability may be largely determined by OB activity that arrives early in a sniff. In this next section, we highlight the importance of measurement alignment: where samples fall relative to stimulus onset and relative to any emergent rhythms, and how inhibition delivered through various connectivity can shape which MCs are read out first.

To assess the impact of inhibition on sniff-aligned measurement windows, we simulate single, global-, and stimulus-based connectivity in response to Gaussian mixtures and skewed Gaussians, and impose a simulated "inhalation." First, we impose a step function that represents the stimulus suddenly being turned on and off at regular intervals. Then, we use a more realistic sniff shape constructed using the difference of exponential functions.

### 4.0.1. On/off step function

As a first step towards implementing a realistic sniff cycle, we divide our time steps into 500 ms-intervals. In each interval, stimulus is presented at time 50 ms and shut off at 450 ms (Figure 4.1). The system is evolved for 200,000 ms, or 400 sniffs in total, and samples for discriminability are taken across sniffs. Network parameters are the same as in steady-state: inhibitory coupling $w$ takes values from 0 to 50 (arbitrary units). We choose measurement durations of size $T = 2, 5, 10, 20$, and 50, to represent sampling at different frequencies. Rather than place measurement windows randomly in time, we align them to sniff onset and slide them along the sniff every 2 ms increments.

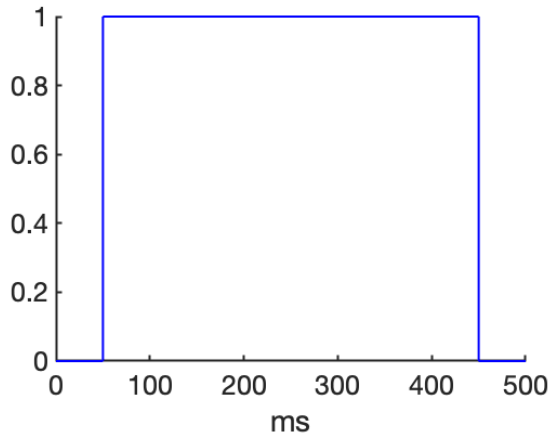


Figure 4.1. Relative stimulus intensity over a 500 ms sniff cycle.

In the interest of reducing visual clutter, we include results from a few select parameters. See Appendix A.10 for a full gamut of activity heatmaps and discriminability plots.

Results from steady-state single- and globally-connected networks, in which measurement windows were taken randomly across the time series, are consistent with trends seen

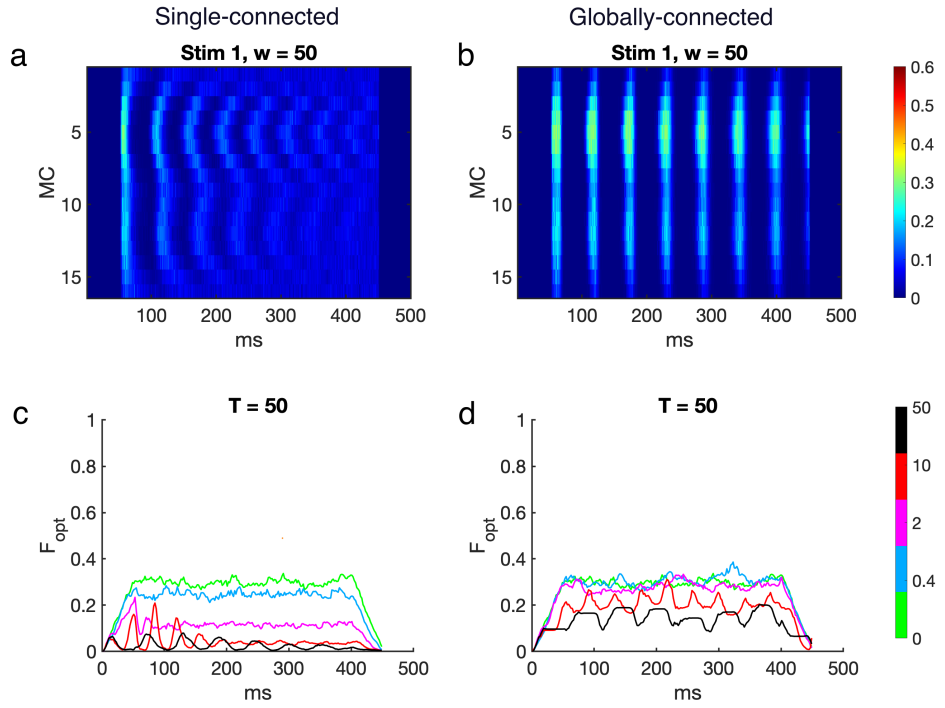


Figure 4.2. Single- and globally-connected networks exposed to Gaussian mixtures as a step function in time. (a) Average firing rates at $w = 50$ in a single-connected network. (b) Same as (a) for globally-connected network. (c) $\mathcal{F}_{opt}$ scaled with $T$ in single-connected network. Samples are of size $T = 50$ and different colored lines from green to black indicate increasing $w$. (d) Same as (c) but for a globally-connected network.

in the latter portions of sniff cycles. Discriminability worsens with inhibition in a single-connection network and trends downwards with inhibition in a global network, especially at larger $w$ (Figure 4.2).

Inhibition given through global connectivity in particular generates rhythmic bands of activity consisting of short volleys of MC spiking alternating with periods of inhibitory-driven silence (Figure 4.3ab). During these volleys, only MCs that are most strongly driven are the most active, and these strongly-driven MCs dominate the system response.

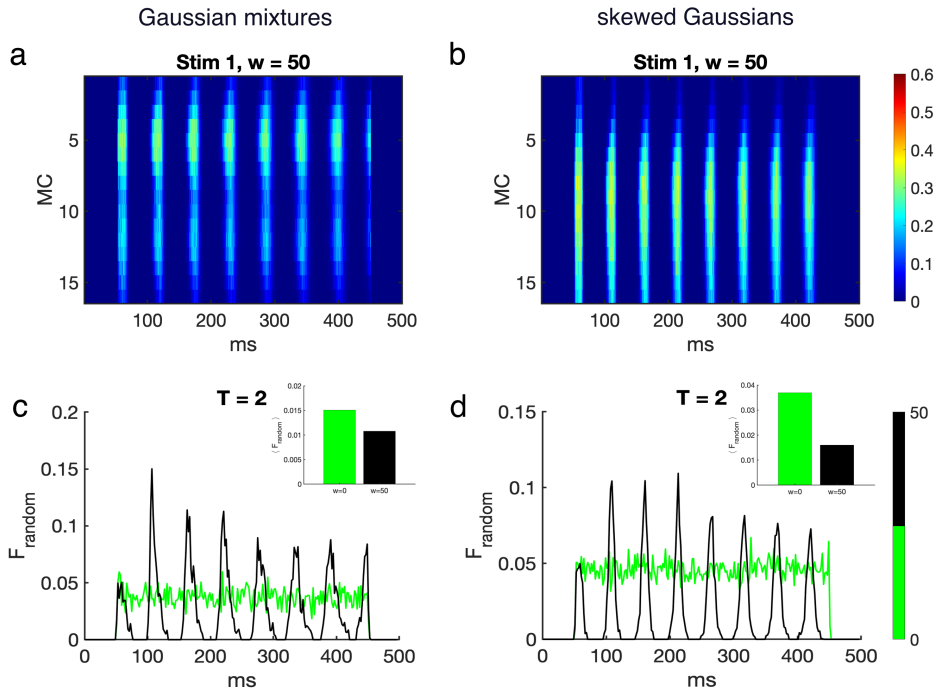


Figure 4.3. Top row: Average sniff of globally-connected network exposed to (a) Gaussian mixtures and (b) skewed Gaussians. Bottom row: Discriminability taken with measurement duration $T$ = 2 ms. Green line indicates $w$ = 0 and black line indicates $w$ = 50. Data shown is $\mathcal{F}_{random}$ scaled with $T$. Insets of (cd) show average over the sniff.

Short measurement windows pick up these highly-active cells. For Gaussian mixtures, discriminability is strongly enhanced in these short intervals, especially when its primacy set (MCs in the centers of the peaks) is active (Figure 4.3c). For skewed Gaussians, discriminability is strongly enhanced when cells in the wings are more active (Figure 4.3d). In both cases, inhibition can greatly improve discriminability in short intervals even though its long-term average (what was measured in the steady-state regime) may be reduced (Figure 4.3cd, insets).

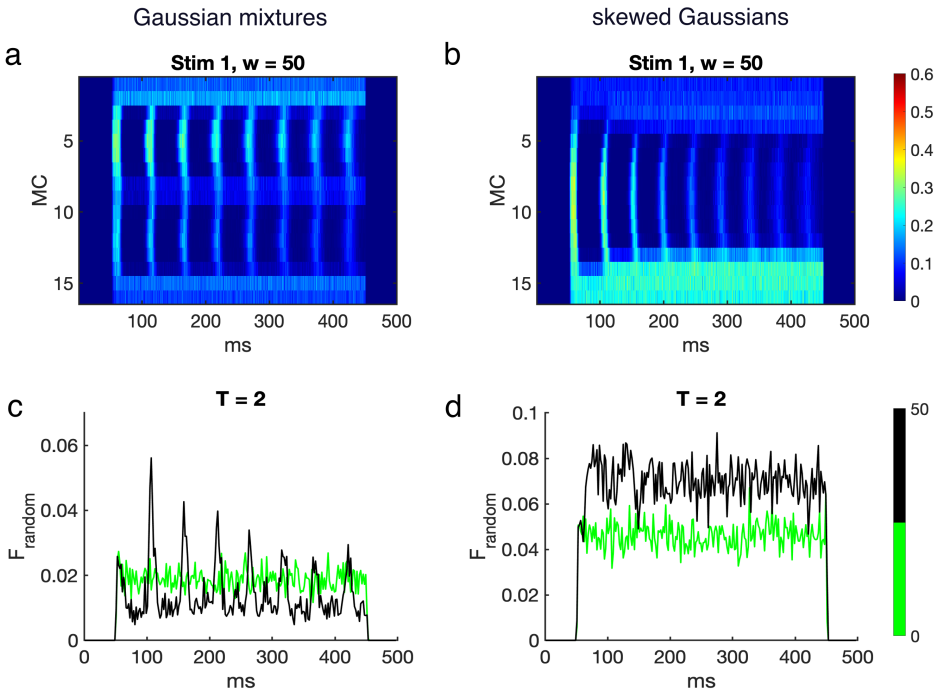


Figure 4.4. Top row: Average sniff of network with stimulus-based inhibition, exposed to (a) Gaussian mixtures and (b) skewed Gaussians. Bottom row: Discriminability taken with measurement duration $T = 2$ ms. Green line indicates $w = 0$ and black line indicates $w = 50$. Data shown is $\mathcal{F}_{random}$ scaled with $T$.

For Gaussian mixtures, stimulus-based inhibition has a similar effect on $\mathcal{F}_{random}$ as global connectivity. Discriminability is improved in small intervals corresponding to activities of the primacy set (Figure 4.4c). These short periods of improved discriminability emerge only when measurement windows are taken with respect to inhalation, and were not observed when measurement windows are taken randomly in time as in the steady-state analyses.

Stimulus-based inhibition has the greatest impact on skewed Gaussians. In that scenario, inhibition promotes sustained improvement of random discriminability by way of suppressing the middle cells whose activity is irrelevant for discriminability (Figure 4.4d).

In doing so, it ensures that OB activity is comprised almost entirely of spikes from cells in the wings, which exhibit the greatest difference between stimuli, and $\mathcal{F}_{random}$ is larger with inhibition across the entire sniff.

### 4.0.2. More realistic sniff cycle

Finally, we consider an inhalation implementation that better matches experimental measurements of internal nasal pressure [**38, 35, 39**]. At each iteration, all OSN rates are multiplied by the value of a double exponential function in time, shown in Figure 5.2. Stimulus intensity ramps up quickly before tapering more slowly to a reset each 500 ms. Equations and parameters are listed in Section 5.

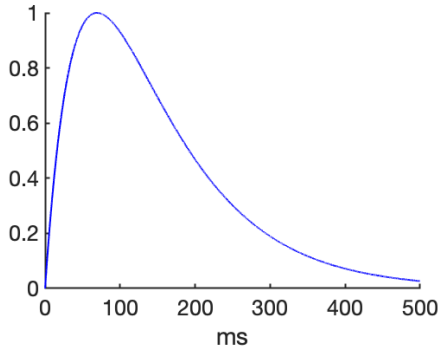


Figure 4.5. Relative stimulus intensity over a 500 ms sniff cycle.

When the sniff intensity is nonconstant, network rhythms are less distinct and vary in length and duration. There are no longer clear visual bands but rather muddier stripes that lose definition over time, resulting from the variability in MC-GCs interaction times and in stimulus intensity itself. Nevertheless, stimulus-based inhibition still improves random discriminability (Figure 4.6) over several gamma cycles.

Taken together, aligning measurement windows to sniff onset allows for better understanding of how information evolves over the course of a sniff. Recurrent inhibition generates patterns of alternating inhibition with activity, which can be either detrimental or beneficial depending on stimulus profile and connectivity regime. Inhibition improves discrimination of Gaussian mixtures specifically at times when highly-excited cells at the

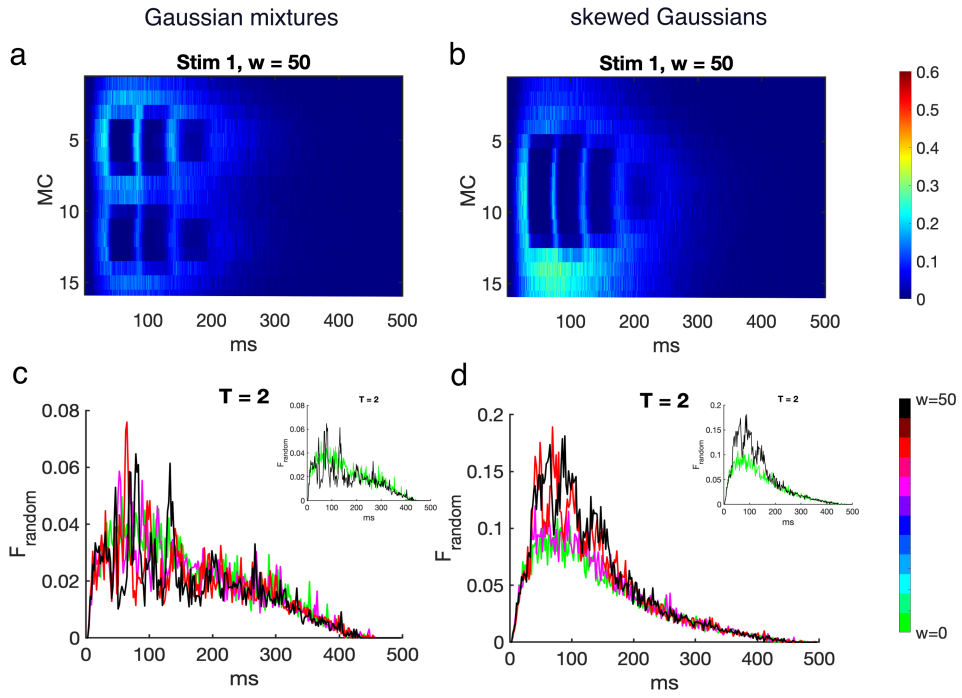


Figure 4.6. Top row: Average sniffs of networks with stimulus-based connectivity exposed to (a) Gaussian mixtures and (b) skewed Gaussians, for a 500 ms sniff cycle. Bottom row: Discriminability taken with measurement duration $T$ = 2 ms. Green line indicates $w$ = 0 and black line indicates $w$ = 50. Data shown is $\mathcal{F}_{random}$ scaled with $T$. Differently colored lines from green to black correspond to increasing inhibition. Insets show only data corresponding to $w$ = 0 and $w$ = 50.

peaks are active, whereas skewed Gaussians are most discriminable when cells in the wings are active. Inhibition induces network rhythms that permit these respective windows to happen, and improvements persist (albeit with less visual regularity) even if sniff intensity is not constant in time.

As in the steady-state, stimulus-based inhibition promotes sustained improvement of exploratory discriminability. These data suggest that networks can better discriminate if they are able to suppress distracting information, especially if that information is the first to arrive in the OB.

CHAPTER 5

# Implementation summary

## 5.1. Neurons are modeled as exponential integrate and fire neurons.

All mitral and granule cells are modeled as exponential integrate and fire neurons. When they reach a fixed threshold potential $V_T$, an action potential is recorded, and the neuron's voltage is set to resting potential.

$$C\frac{dv}{dt} = -g_L\left(v - E_L\right) + g_L\Delta_T e^{\frac{v-v_T}{\Delta_T}} + I \tag{5.1}$$

$$v > V_T \rightarrow v = E_L. \tag{5.2}$$

Here, $C$ is the membrane capacitance, $g_L$is the leak conductance, $E_L$ is the resting potential, $\Delta_T$ is a slope factor, and $I$ represents input current. Olfactory sensory neurons are modeled as Poisson spike trains with fixed mean rates over time, which represents how the system may behave under steady-state conditions when a chemical stimulus is continually available. Each stimulus "shape" is a mapping of OSNs to average rates; there is otherwise no noise input into the model.

Each GC, MC, and OSN has a low-pass filter variable $\bar{x}$ that is incremented following an action potential and decays exponentially with time constant $\tau_x$.

$$\tau_x\frac{d\bar{x}}{dt} = -\bar{x} \tag{5.3}$$

$$v > V_T \rightarrow \bar{x} = \bar{x} + 1 \tag{5.4}$$

Interactions between MCs and GCs are governed by two matrices: $W_{mg}$ and $W_{gm}$ that represent the connectivities from MCs to GCs and GCs to MCs, respectively. These naming conventions ostensibly go against traditional naming rules; the first subscript indicates the source and the second indicates the recipient. There is a third diagonal matrix $W_{om}$ that represents the single one-way excitatory connections from OSNs to MCs

$$
\begin{aligned}
\tau_{\mathrm{MC}_k}\frac{d}{dt}\bar{x}_{\mathrm{MC}_k} &= -\bar{x}_{\mathrm{MC}_k} + \sum_{\mathrm{MC}_k \text{ spike}} \delta(t - t_{\text{spike}}) \\
\tau_{\mathrm{GC}_i}\frac{d}{dt}\bar{x}_{\mathrm{GC}_i} &= -\bar{x}_{\mathrm{GC}\ i} + \sum_{\mathrm{GC}_i \text{ spike}} \delta(t - t_{\text{spike}}) \\
I_{\text{to MC}} &= -\left(\sum_i w \times W_{gm} \times \bar{x}_{\mathrm{GC}\ i}\right) + \left(\sum_j W_{om} \times \bar{x}_{\mathrm{OSN}\ j}\right) \\
I_{\text{to GC}} &= \left(\sum_k W_{mg} \times \bar{x}_{\mathrm{MC}\ k}\right),
\end{aligned}
\tag{5.5}
$$

where $i$ indexes all granule cells, $j$ indexes OSNs, $k$ indexes mitral cells, and $w$ is a nonnegative scalar representing inhibitory coupling strength.

In each steady-state analysis, the network is stimulated for 50,000 milliseconds using a Forward Euler time stepping scheme with $dt = .02$. Measurement windows of a fixed duration are randomly placed throughout the duration of the time series, and we calculate linear discriminability out of those samples. We allow for the possibility that measurement windows overlap and are therefore not independent of each other; however, we find that probabilities of overlapping windows do not affect numerical estimates of $\mathcal{F}_{opt}$ (see Appendix A.2).

## 5.2. Linear discriminability

Linear discriminability measures the discriminability or dissimilarity of two network representations of stimuli.

### 5.2.1. Optimal

Optimal Fisher information is the maximum noise to signal ratio achieved by projecting multidimensional inputs onto a line in multidimensional space.

$$\mathcal{F}_{opt} = \max_{\vec{c}} \frac{\left(\vec{c} \cdot \langle \vec{M}^{(1)} \rangle - \vec{c} \cdot \langle \vec{M}^{(2)} \rangle\right)^2}{\vec{c}^T \left(\sum^{(1)} + \sum^{(2)}\right) \vec{c}}, \tag{5.6}$$

where $\vec{c}$ is a vector of weights, $\langle \vec{M}^{(i)} \rangle$ is a length-$N$ vector containing the average responses to stimulus $i$, and $\Sigma^{(i)}$ is the corresponding covariance matrix. Note that the magnitude of $\vec{c}$ is divided out; $\mathcal{F}_{opt}$ only depends on its direction. In this thesis, we report $\mathcal{F}_{opt}$ scaled by $T$ to account for known improvement with increasing measurement bin.

For ease of notation, we denote $\Delta\vec{\mu}$ as the difference in average responses an $\Sigma$ as the sum of covariances, respectively:

$$\Delta\vec{\mu} = \langle \vec{M}^{(1)} \rangle - \langle \vec{M}^{(2)} \rangle \tag{5.7}$$

$$\Sigma = \Sigma^{(1)} + \Sigma^{(2)}, \tag{5.8}$$

allowing us to write $\mathcal{F}_{opt}$ as the optimal ratio of the squared projection of $\Delta\vec{\mu}$ to the projection of noise terms $\Sigma$:

$$\mathcal{F}_{opt} = \frac{\left(\vec{c}_{opt} \cdot \Delta\vec{\mu}\right)^2}{\vec{c}_{opt}^T \Sigma \vec{c}_{opt}}. \tag{5.9}$$

It can be shown that the optimal vector of weights $\vec{c}$ is proportional to $\Sigma^{-1}\Delta\vec{\mu}$, and we can simplify optimal linear discrimination as follows:

$$\mathcal{F}_{opt} = \frac{\left(\Sigma^{-1}\Delta\vec{\mu}\cdot\Delta\vec{\mu}\right)^2}{\left(\Sigma^{-1}\Delta\vec{\mu}\right)^T\Sigma\left(\Sigma^{-1}\Delta\vec{\mu}\right)} \tag{5.10}$$

$$= \Delta\vec{\mu}^T\Sigma^{-1}\Delta\vec{\mu}. \tag{5.11}$$

Therefore, $\mathcal{F}_{opt}$ is proportional to the squared difference in mean activities and inversely proportional to the sum of covariances.

In this work, we often compare $\mathcal{F}_{opt}$ with $\mathcal{F}_{uncorr}$, an equivalent computation that does not include any pairwise correlations. That is,

$$\mathcal{F}_{uncorr} = \Delta\vec{\mu}^T\Sigma_{diag}^{-1}\Delta\vec{\mu}, \tag{5.12}$$

where $\Sigma_{diag}$ has the same entries as $\Sigma$ on the main diagonal and is zero everywhere else.

**5.2.1.1. Decomposition of $\mathcal{F}_{opt}$ with respect to $\sum$.** It is sometimes useful to decompose $\mathcal{F}_{opt}$ into the sum of contributions corresponding to each eigenmode of $\Sigma$.

$$\mathcal{F}_{opt} = \left(\sum_j\left(\Delta\vec{\mu}\cdot\vec{e}_j\right)\vec{e}_j\right)^T\Sigma^{-1}\left(\sum_k\left(\Delta\vec{\mu}\cdot\vec{e}_k\right)\vec{e}_k\right) \tag{5.13}$$

$$= \left(\sum_j\left(\Delta\vec{\mu}\cdot\vec{e}_j\right)\vec{e}_j\right)^T\left(\sum_k\frac{1}{\lambda_k}\left(\Delta\vec{\mu}\cdot\vec{e}_k\right)\vec{e}_k\right) \tag{5.14}$$

$$= \sum_j\frac{1}{\lambda_j}\left(\Delta\vec{\mu}\cdot\vec{e}_j\right)^2. \tag{5.15}$$

### 5.2.2. Random / exploratory

$\mathcal{F}_{opt}$ and $\mathcal{F}_{uncorr}$ both measure separability under optimal conditions; these involve projecting two stimuli's means and covariances onto a separatrix using weights that maximize the signal to noise ratio. However, we also consider discriminability that is more akin to random guessing, for instance to represent when an animal is first exposed to a set of unfamiliar stimuli. Such a metric uses random weights instead and assumes no prior knowledge of how to weight different inputs. Visually, this entails drawing a line through the weighted center of an inputs in a random direction, in which the probability of each line is rotationally invariant. Several examples of such a separating line are shown in Figure 5.1.

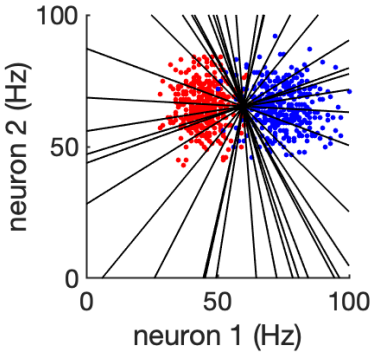


Figure 5.1. Two clouds with rotationally invariant separating lines. In this schematic, weight vectors (and correspondingly, separating lines) are drawn through the weighted center of the two stimuli.

We denote this "random discriminability" as $\mathcal{F}_{random}$ and compute it as follows.

$$\mathcal{F}_{random} = \frac{\left(\vec{c}_{rand} \cdot \left(\langle \vec{M}_1 \rangle - \langle \vec{M}_2 \rangle\right)\right)^2}{\vec{c}_{rand}^{\,T} \left(\Sigma_1 + \Sigma_2\right) \vec{c}_{rand}}, \qquad \vec{c}_{rand} = \left[x_1, ..., x_N\right]^T, \qquad x_i \sim iid\ \mathcal{N}(0,1).$$

To easily see why this sampling gives us a rotationally invariant weight vector, consider the product of probability distributions $\Pi_i^N p(x_i)$.

$$p(x_1, ..., x_N) = \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_1^2}\right)\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_2^2}\right) ... \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_N^2}\right) \tag{5.16}$$

$$= \left(\frac{1}{(2\pi)^{N/2}} e^{-\frac{1}{2}\left(\sqrt{x_1^2+x_2^2+...+x_N^2}\right)^2}\right), \tag{5.17}$$

which is only a function of the magnitude of $\vec{c}_{rand}$ and not of its directional orientation.

Remark: The algorithm produces iterations of vector $\vec{c}_{rand}$, whose components are drawn i.i.d. from a standard normal distribution. The resulting vector points outward from the origin. The diagram shown in the future above however draws these rotationally invariant lines outward from the weighted center of the two activity clouds. Intuitively, translating the two activity clouds so that their centroid sits on the origin would not affect the separability of two clouds. We also show this analytically in Appendix B.2.

### 5.2.3. Comparison to artificial Poisson spike trains

We often plot networks' performance against those of Poisson spiking networks with the same mean rates. These "networks" suppose individual neurons are independent from each other, and each neuron's variance of spike counts is equal to its mean. It follows that the covariance matrix $\Sigma^{(i)}$ is a diagonal matrix whose main diagonal is equal to $\langle \vec{M}^{(i)} \rangle$. Under these assumptions, $\mathcal{F}_{opt}$ takes the form

$$\mathcal{F}_{opt} = \sum_{MC\ i} \frac{\left(M_i^{(1)} - M_i^{(2)}\right)^2}{M_i^{(1)} + M_i^{(2)}}, \tag{5.18}$$

where $M_i^{(j)}$ is the response of mitral cell $i$ to stimulus $j$. Note that in general we cannot assume any relationships between $\Delta\vec{\mu}$ and $\Sigma$. However, Poisson spiking properties allow us to compare network data against a baseline case where the variance of rates has a known reduction with respect to reduction in means.

### 5.3. Timing-based metrics and the sniff cycle

In Section 4, we simulate the sniff cycle in two ways (Figure 5.2). The first is with a Heaviside function, which in each 500 ms interval, turns on at 50 ms and turns off at 450 ms. Networks are evolved for 200,000 ms, or 400 "sniffs" in total. At each time increment, discriminability is computed using samples taken randomly from a proper subset of the 400 sniffs.

The second implementation uses the difference of two exponentials to mimic nasal pressure over a sniff. At each timestep, all OSN rates are multiplied by the value of a double exponential function in time. Equations are shown for the 2 Hz, which is characterized by stimulus intensity $S$ ramping up rapidly, followed by tapering down more slowly, before resetting.

$$S(t) = S_{\max}\left[\exp\left(\frac{t_{mod}}{\tau_1}\right) - \exp\left(\frac{t_{mod}}{\tau_2}\right)\right] \tag{5.19}$$

$$\text{where } t_{mod} = (t \mod 500),\ \tau_1 = 50,\ \text{and } \tau_2 = 100. \tag{5.20}$$

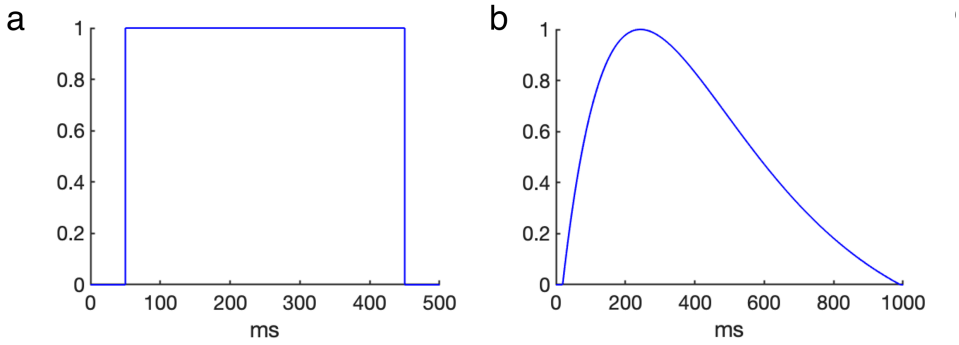


Figure 5.2. (a) Step function (b) Simulated inhalation.

CHAPTER 6

# Conclusion and Discussion

This work explains how several physiological features of the olfactory bulb (OB) contribute to sensory discrimination. Underlying our results are the gamma oscillations observed in OB, driven by the reciprocal synapses between excitatory mitral cells and GABAergic granule cells [**40, 41, 42, 43**]. These inhibitory granule cells reformat OB activity representations within the OB over the course of a sniff, reducing the pairwise correlation between responses to different stimuli [**7, 8**]. In doing so, they reshape mitral cell responses before they are read out by downstream piriform cortical (PCx) neurons, which integrate OB activity in order to enable odor recognition and classification. In general, while there is always a tradeoff between decision speed and accuracy (a longer readout would always enable PCx to gather more information), the dominant response from higher-order neurons occurs within tens of milliseconds after odor onset [**37, 44, 45, 46**] suggesting that OB information necessary for stimulus discrimination must be delivered within a finite-duration window shortly after inhalation. In this thesis, we have first shown in the steady-state that noise correlations between MCs vary on the gamma timescale imposed by recurrent inhibition, which enhances linear discriminability in short measurement durations when the stimulus pair is low-contrast. We have next shown that in a timing-based framework, inhibition improves discriminability shortly after stimulus presentation, at times during which differentially-excited MCs are most active.

We quantify the separation between two stimuli using linear discriminability, which increases with larger differences in average responses and decreases with increasing noise. Many prior studies measure OB activity using a firing rate framework, which inherently ignores spiking pattern information in favor of a long-term average over some long but finite duration. To understand how inhibition can improve stimulus discriminability, our work first similarly uses a steady-state regime in which stimulus is continually presented and measurement windows of variable size are taken randomly in time. Under this framework, we find that weak inhibition reduces the variance of MC rates (calculated as spike count divided by the finite measurement duration) more strongly than it would in a Poisson neuron, whose variance scales linearly with average spike count. This variance reduction is stronger in smaller interconnected circuits, in which GC activity is more correlated with that of individual MCs and can more effectively reduce their noise.

Moreover, increased inhibitory coupling drives the gamma rhythm to have higher amplitude and lower frequency, and emergent noise correlations adhere to the oscillations along which MCs and GCs spike. The variance (scaled by measurement duration) of individual MCs' finite-time rates is lowest at measurement durations near that system period, and linear discriminability is accordingly highest in those measurement durations, even after accounting for an expected improvement with larger readout time. In other words, recurrent inhibition concurrently shapes MC response in two ways; it induces oscillations in stimulus-driven activity response, and it reduces MCs' variance, improving discriminability in short finite-window durations near the gamma period.

Shared connections also induces pairwise covariances between different MCs, and following various studies in the literature [**21, 15, 17, 18**], we assess whether or not these

pairwise noise terms increase information needed for discrimination. In one of these studies, the authors find that lateral connections between cells reduce the total amount of information available, and that the majority of remaining information lies in the pairwise covariance terms [**16**]. While the lateral connections in their study specifically mimic the local-excitation-lateral-inhibition properties of the visual cortex, we only consider noise-shaping properties of inhibitory lateral connections and show that the contributions of noise correlations additionally depend on the stimulus contrast and measurement duration. Furthermore, whether or not noise correlations improve discriminability is more complex and depends on the relative magnitudes between the covariance of two similarly-stimulated MCs and that of two differently-stimulated MCs across both stimuli in a classification task. When stimuli are low-contrast, the overall contribution of covariance terms is beneficial towards discriminability, whereas when stimuli are high-contrast, they are detrimental. There are studies that suggest that OB activity responses are reformatted differently depending on task difficulty [**6**]. and our work provides a concrete mathematical suggestion for different modalities: a network faced with a low-contrast discrimination task would benefit from using information from pairwise covariance terms, whereas a network facing a high-contrast pair of stimuli would be better off without it.

This work is not the first effort to explain how noise correlations and discriminability vary on different timescales in the olfactory system. Giridhar et al. 2011 find that inhibition-induced pairwise correlation between MCs is positive at short timescales and negative at long timescales. The former reflects MCs' synchronizing on a gamma timescale, enhancing the likelihood that spikes propagate to the cortex at the cost of

information quantity, whereas the latter corresponds to more total information and enhances encoding potential [**23**]. We supplement this finding of correlations' changing sign by showing how pairwise covariances change sign precisely at multiples of the system period, where the period itself is dictated by the frequency of recurrent inhibition-driven gamma oscillations. Indeed, small measurement windows capture synchronous firing from MCs spiking in the same volley, whereas at larger windows, some highly-excited MCs spike more frequently, indirectly silencing its weakly-stimulated counterparts. Additionally, where the authors of Giridhar et al. quantify correlations using Pearson correlation coefficient (inherently normalized) relative to that of MCs without shared connections, our usage of variance and covariance allows us to directly calculate their effects on discriminability as well as isolate their contributions from that of mean difference.

In the field of sensory processing, evolution of synaptic weights in response to experience is widely considered and studied [**32, 33, 6, 34**]. We implement this idea by assessing networks whose connectivity could be the result from exposure-driven learning: MCs who receive more stimulation are more strongly coupled to their GCs, also indirectly inhibiting themselves more strongly. As a result, these stimulus-shaped networks preferentially silence MCs who are most strongly stimulated, allowing MCs that are more weakly stimulated to dominate the network response. This enhances discriminability during tasks in which the differences between a pair of stimuli lie in more weakly-excited MCs and when the readout is exploratory and has not been already optimized to favor differentially-excited MCs. In short, a network that has been shaped by experience-dependent learning can perform better discrimination by preferentially silencing highly-stimulated MCs whose activity contributes to total noise but does not contribute to average mean difference.

A large majority of our analyses use a steady-state framework to understand how noise correlations can enhance discriminability in small, finite-time measurement durations dictated by network gamma oscillations. In the last chapter of this thesis, we extend our analyses to a timing-based framework. An inhalation cycle is modeled by "turning on" the stimulation at regular intervals and turning it off at the end of each simulated sniff cycle, first with a Heaviside step function, then with a double exponential function that mimics nasal pressure during inhalation. These efforts are motivated by the results of some recent works in olfactory processing that find information may be encoded in a few early spikes from MCs, rather than in long-term average firing rates [**31**]. In line with these studies that show the importance of early encoding, we similarly show that inhibition can enhance discriminability shortly after stimulus is first presented to the olfactory system.

Gamma oscillations that were observed in steady-state also persist in the sniff framework, and in order to assess how the readout changes over the course of inhalation, we align measurement windows to sniff onset rather than randomly in time. Similarly as in steady-state, we find that inhibition enhances discriminability in short, finite-time measurements and maximal discrimination occurs when MCs that are most differentially-excited are spiking, regardless of which MCs contribute to stimulus difference. If the most active MCs are differently excited across two stimuli, then discriminability is highest in early gamma cycles during which those MCs spike for the first time. Otherwise, the network is better off waiting for a later gamma cycle, after which recurrent inhibition silences the strongly-excited MCs, allowing more weakly-stimulated MCs to be read out. This latter result is especially pronounced when the connectivity matrix is "learned," and

early-spiking MCs that do not contribute to discriminability are quickly suppressed to make way for more important OB information. Thus, inhibition can improve stimulus discrimination shortly after inhalation begins regardless of stimulus profile.

There are various olfactory tasks and situations that are not explicitly considered but that our modeling analyses can help understand. For instance, this thesis only considers pairs of stimuli that are symmetric, representing binary mixtures commonly used in experimental learning tasks [**27, 28, 29**]. We may additionally consider pairs of stimuli of which one is a scalar multiple of the other, for instance representing different concentrations of the same chemical substance. In this case, positive covariances between similarly-stimulated MCs would be beneficial, rather than harmful, for discriminability. Another assumption our model makes is that sniffs are always 2 Hz; in practice, the durations and intensities of firing activity over a sniff cycle may vary with behavior (exploration vs. learning) or with the animal's wake state. Considering how odor representations change with behavior would be an interesting extension of our current work but not directly relevant in showing how inhibition-induced gamma oscillations shape noise correlations and discriminability in a time-dependent manner.

In the appendix, we discuss other ways in which this model can be further analyzed. While the networks listed in the main text contain at most 16 MC, we find that inhibition improves discrimination more weakly at larger system sizes and that greater numbers of interconnected MCs dilute the noise-shaping – and in particular the variance-reducing – effects of recurrent connectivity (Appendix B.4.4). Additionally, the effect of inhibition on discriminability is largely impacted by the difference in average stimulus response $\Delta\vec{\mu}$, which depends heavily on the neurons' sensitivity to input and thus the choice of

neuron model (adaptive integrate and fire, stochastically firing neuron, etc.) or which properties to include (e.g. refractory period). Future studies may explore how variances and $\Delta\vec{\mu}$ shape discriminability using these alternative neuron models, which are discussed in Appendix B.6 and B.7.

Another modeling choice we made was to implement stimulus-based "learned" connectivity by modifying the mitral-to-granule coupling matrix $W_{mg}$ rather than $W_{gm}$. We found in the latter scenario that the resulting network behaves similarly, as long as the overall inhibition delivered to MCs was similar. However, this consistency relies on GCs in the model not receiving external noise. In appendix B.4, we show via a linear rate model that MC noise is amplified if GCs receive external noise, which implies that altering $W_{gm}$ instead of $W_{mg}$ may affect how recurrent inhibition shapes noise correlations.

This work quantifies odor dissimilarity using linear discriminability, in which MC input is weighted akin to how the cortex may weight olfactory inputs in formulating a final decision. There are however many alternative metrics that may be used in further explorations. Linear discriminability measures the dissimilarity of odor representations, which tells us something about how difficult a task is. Other studies quantify dissimilarity by framing odor discrimination using a probabilistic approach [**47**], as a set theory problem [**48**], or using a diffusion-decision model that demonstrates how the output of a linear classifier evolves over time [**49**]. In linear discriminability, $\mathcal{F}_{opt}$ also directly correlates with the percentage of total samples correctly classified under the optimal separatrix; some examples are included in appendix B.4.4. This thesis focuses on $\mathcal{F}_{opt}$ specifically, a numerical score that allows us to directly mathematically quantify how noise correlations and $\Delta\vec{\mu}$ individually impact discriminability.

In summary, this work shows how the gamma oscillations within the olfactory bulb dictate how noise correlations between MCs fluctuate, shaping discriminability on the same timescale. It also shows how inhibition reduces variance of finite-time rates and induces pairwise covariances between MCs, which enhance discriminability when stimuli are low-contrast. In a simulated sniff cycle, recurrent inhibition shapes MCs to spike in rhythmic volleys inside which MCs that are relevant for discrimination can be read out. Taken together, this work demonstrates the mechanistic importance of shared inhibitory connections in the bulb and quantifies precisely how spiking statistics shaped by these recurrent interactions enhance pattern separation on a cortically relevant timescale.

CHAPTER 7

# Mechanistic modeling in clinical pharmacology

In summer 2021 I completed a summer internship with Takeda Pharmaceuticals and learned how to use mathematical modeling techniques to inform drug development and clinical dosing strategies. In this next section, I describe my internship project: development of a mechanistic model to understand preclinical-to-clinical translation of CAR-T cell therapies.

Despite groundbreaking clinical success of CAR-T cell therapies, a translational gap exists between preclinical and clinical pharmacology, and interspecies differences in dose-exposure-response relationships are poorly understood. To address this gap in translational research, we selected a CD19 directed 2nd generation CAR-T construct [**50**] and developed a unified physiology-based pharmacokinetic-pharmacodynamic (PBPK-PD) model that characterizes CAR-T and B-cell dynamics in both mice and patient populations simultaneously. This model addresses CAR-T cell distribution in peripheral blood (PB) and other pertinent tissues ($n$ = 20) in NCG immunodeficient mice both with and without disseminated CD19+ NALM-6 cells, and cellular kinetics and B-cell depletion in 13 cancer patients. Key model components include both nonspecific and target-engagement-mediated CAR-T expansion, distribution of CD19+ B-cells in lymphoid organs, and CAR-target engagement kinetics. Many of the estimated parameters, which include first-order tissue transmigration rate constants, first-order maximum killing

rate constants, and first-order maximum expansion rate constants, suggest key mechanistic differences in CAR-T cell expansion and efficacy between mice and humans. Using this case study, we can propose inter-species translational rules on estimated expansion and efficacy parameters that could provide insight towards preclinical-to-clinical translation of CAR-T cell therapy treatments.

## 7.1. Introduction

Chimeric antigen receptor (CAR) T-cell therapies comprise a family of immunotherapeutic treatments that have enjoyed revolutionary clinical success in treating leukemias, lymphomas, and myelomas [**51, 52, 53, 54, 55, 56**]. They have been described as a miracle drug for many patients with relapsed or refractory cancers, in some cases eradicating cancer cells up to a decade after treatment. Their successful clinical outcomes have spurred rapid development and approval of marketable drugs – in the last several years alone, five CAR-T cell products have been approved by the United States Food and Drug Administration (Abecma®, Kymriah®, Breyanzi®, Tecartus® and Yescarta®) and one product (Relma-cel®) has been approved in China [**57, 58, 59, 60**].

Despite its success however, CAR-T cell therapy is usually a last-resort for cancer patients and is only prescribed after other treatments have failed. Indeed, CAR-T cell therapies are often studied in patients with relapsed or refractory cancers who have exhausted other options and have low life expectancy. Even at that stage, CAR-T treatment is not only also expensive – according to Novartis, the manufacturer of Kymriah, a single-treatment course of Kymriah costs $475,000 – but also risky and high-stakes [**61**]. To that

end, gaining a better mechanical understanding of the pharmacokinetic and pharmacodynamic properties of CAR-T cell therapies is imperative toward advising clinical dosing decisions. In this project, we develop a translational model that describes the cellular kinetics and biodistribution of a CAR-T construct in both murine and human populations, with the longer term goal of furthering our collective understanding of interspecies differences in cell therapy dose-exposure-response.

The mechanisms of CAR-T cell therapies involve isolating T-cells from cancer patients and genetically engineering them to express target-specific receptors. These chimeric antigen receptors (CARs) are designed to specifically recognize cancer cells upon reinfusion into the patient, with the goal of seeking out and killing tumorous cells that express those specific antigens. After infusion, CAR-T cells exhibit a distinctive profile that is characterized by four phases: distribution, expansion, contraction, and persistence, with various disease-related and drug-specific characteristics known to influence the shapes and durations of each phase [**62, 63**]. We focus specifically on species-specific contributions to CAR-T proliferation and killing efficiency in order to better understand how dose-exposure-response relationships scale between preclinical models and clinical trials of the same product.

Many previous studies use PBPK-PD models to estimate parameters related to biodistribution of various biologics and cell types, including monoclonal antibodies, T cells, and natural killer (NK) cells [**64, 65, 66**]. In recent years, many of these models have been adapted for CAR-T studies and estimate distribution and expansion parameters, such as tissue-specific first-order transmigration rates, first-order CAR-T metabolic rate constants, target-mediated killing and expansion rates [**67, 68**]. Our work builds upon these

existing efforts in a few key ways. First, while both models use compartmentalized representations of major body systems, we expand the structural model to include significantly more tissue types (Figure 7.1a), allowing for highly detailed analyses of biodistribution dynamic, as well as laying the groundwork for more fine-tuned covariate analyses. Another key difference lies in how tumorous target cells are represented in the model. Since this specific CAR-T construct targets CD19+ B cells, we represent their time-evolution dynamics to include the B-cell specific processes of maturation, apoptosis, generation, and trafficking between specific tissues (Figure 7.1b), rather than implementing passive diffusion within the tumor compartment like some previous works have done. These structural components of B cell dynamics are also more compatible with the available experimental procedure, which comprises intravenous injection of NALM-6 cells that then diffuse through the peripheral blood and distribute through the extravascular compartments of select tissues (Figure 7.1c). We acknowledge that this assumption may not fully account for target-engagement mechanics in studies of solid tumors; however, it more accurately represents of the true biological natures of haematological malignancies, in which cancerous cells diffuse through the peripheral blood and lymphatic organs. More precise modeling of CAR-T interactions with solid tumors are outside the scope of this work and are the subject of future modeling efforts. Taken together, our work provides a comprehensive, increased biophysically accurate modeling framework that simultaneously characterizes murine and human biological dynamics at an unprecedented level of intricacy, and the biophysical parameters estimated by the model suggest interspecies differences in CAR-T pharmacokinetic and pharmacodynamic properties. These modeling

efforts also may inform future research in studying preclinical-to-clinical translation of other CAR-T products as well as other cell therapy treatments.

Taken together, our work provides a comprehensive, increased biophysically accurate modeling framework that simultaneously characterizes murine and human biological dynamics at an unprecedented level of intricacy, and the biophysical parameters estimated by the model suggest interspecies differences in CAR-T pharmacokinetic and pharmacodynamic properties. These modeling efforts also may inform future research in studying preclinical-to-clinical translation of other CAR-T products as well as other cell therapy treatments.

## 7.2. Model-fitted profiles

### 7.2.1. A mechanistic model addresses CAR-T biodistribution and nonspecific expansion in NCG immunodeficient mice without CD19+ B cells:

Figure 7.2 shows model-fitted CAR-T concentration profiles (solid lines) against observed data (blue points) in seven tissue types in male (left panel) and female (left panel) NCG immunodeficient mice. Horizontal axis indicates time measured in hours following infusion of $1 \times 10^7$ CAR-T cells, and mice were killed and CAR-T concentration measured at three hours, two days, eight days, and 15 days post infusion. Each experimental datapoint represents the average value over three animals. Data and model fitted profiles are plotted on a log scale.

For all tissues ($n$ = 7) in this cohort, volumes, blood flow rates, and lymph flow rates were fixed to reported literature values. Since these animals have no functional immune system and did not receive any B cell injections, we conclude that cells proliferate

through some nonspecific expansion parametrized by doubling time ($DT$) and maximum blood concentration ($C_{Blood}^{max}$), both of which are estimated by the model. Describing nonspecific expansion using these two parameters reproduces key visual characteristics of expansion patterns reported in vitro [**69**]. All tissues except blood, bone marrow, and MLN are assumed to have both vascular and extravascular compartments through which CAR-Ts distribute. Flow rate from the vascular to extravascular compartment is parametrized by first-order transmigration rate constant $J_{tissue}$; in our model, $J_{Lung}$, $J_{Spleen}$, and $J_{Liver}$ were fixed to literature [**68**] and the rest were estimated. The resulting model fits reasonably capture the overall increasing trends in most tissues, with a straight increasing line indicating expected exponential growth on the experimental time scale.

### 7.2.2. Expanded model with incorporated B cell dynamics describes CAR-T biodistribution and target-engagement-mediated expansion.

Figure 7.3 shows model fitted profiles for CAR-T concentration (solid lines) and experimental datapoints (blue dots) in male (left panel) and female (right panel) NCG immunodeficient mice that have been injected with $1 \times 10^6$ NALM-6 cells five days before intravenous injection of $5 \times 10^6$ CD19 CAR-T cells. Horizontal axis represents time in hours following CAR-T injection, and all data and model fitted profiles are plotted on a log scale. Blue dots plotted at 77 Copies/$\mu$g gDNA (the LLOQ) represent timepoints for which the measured value was below detectable values.

To accurately characterize B cell dispersion and movement, we include evolution dynamics similar to that reported in Hosseini et al. 2018 and developed for modeling by

Kesisoglou and Maddah [**70**], who were interns concurrently in the same program. Specifically, B cells diffuse between peripheral blood, spleen, liver, and bone marrow at rates whose steady state ratios are fixed quantities known as partition coefficients. Within each specific tissue, existing B cells proliferate and experience apoptosis at rates fixed to the literature [**71**]. CAR-T-target engagement occurs in the extravascular compartments of tissues in which both B cells and CAR T-cells exist. In each compartment, we assume that CAR-T-target complexes form at rates proportional to the availability of B cell receptors and both the binding affinity and dissociation rate of the construct. The concentration of complexes modulates the rate at which CAR T-cells further proliferate in those compartments as well as the rate at which target B cells are destroyed. Because NCG immunodeficient mice do not have functional B, T, and NK cells [**72**], the model assumes that the only B cells present in the system are that which were injected and correspondingly, there is no progenitor B cell population in bone marrow. A consequence of these assumptions are that all available B cells are potential targets; in future work, it would straightforward to categorize B cells into separate target and non-target populations. Transmigration rates for bone marrow, liver, lung, spleen were fixed to the estimated values from fitting the model to non-tumor-bearing NCG mice. The model estimated transmigration rates for other tissues.

For most tissue types, the model reasonably captures the visual characteristics of the initial injection and rapid distribution to tissue types, as well as the slower longer-term rise. In general, the model reliably predicts values below the LLOQ when relevant, although it tends to overestimate CAR-T concentration in the sex-specific organs of testes, epididymis, and ovary. We acknowledge that this limitation may be restricted to lack of

precision in their related parameters, as tissue volumes and flow rates for these organs are not well-studied nor documented.

### 7.2.3. A scaled-up PKPD model addresses cellular kinetics and B cell depletion and recovery in clinical patients:

Figure 7.4 shows model-fitted profiles (black lines) of CAR-T cellular kinetics against observed data (blue dots), in 13 clinical patients. Data show concentration of CAR-T cells in peripheral blood versus days following CAR-T infusion. Figure 7.5 shows model fitted profiles (black lines) and observed data (blue dots) of B cell percentage of seven of the thirteen patients. Observed data represent amount of CD19+ cells in peripheral blood as a percentage of baseline values. All patients experienced at least two lines of prior treatment and went through a preconditioning lymphocyte-depleting regimen of 25 $mg/m^2$ of fludarabine and 250 $mg/m^2$ of cyclophosphamide before receiving doses of $10^6$ CAR-T cells per kilogram of body weight. Clinical data differs from preclinical data in two main ways. Firstly, data is limited to measurements that can be made in peripheral blood (cellular kinetics and B cell concentration) whereas preclinical studies reported biodistribution in several different organs. Secondly, the clinical patients in this study underwent a lymphocyte-depleting regimen, with further B cell aphasia induced by CAR-T target-killing, followed by longer-term B cell recovery. In contrast, all mouse models used in this study are NCG immunodeficient and assumed to have no endogenous B cells.

To address the first challenge, we replicate the full structural model in mice and adjust parameters to suit human physiology. When available, tissue volumes, blood flow rates, and lymph flow rates were fixed to experimental values reported in literature. Unavailable

lymph flow rates were estimated as $\frac{1}{5000}$th of the corresponding blood flow rates, consistent with existing practices in PKPD modeling [**68, 73, 67**]. Unavailable tissue volume rates were estimated from mouse tissue volumes using an allometric exponent of 1, and all first-order transmigration rate constants were allometrically scaled with an exponent of −0.25, using the estimated values estimated from tumor-bearing mice data.

B cell dynamics and CAR-T-target engagement dynamics are captured similarly to that of the NCG immunodeficient mice; B cells diffuse through the peripheral blood, spleen, liver, and bone marrow at rates whose steady state ratios relative to each other are fixed. Formation of complexes drives B cell killing and target-mediated CAR-T expansion within the same extravascular compartments as the complexes. We use Hill dynamics to model both how killing rates and expansion rates depend on the concentration of available complexes, and relevant parameters in all Hill equations are estimated by the model.

Unlike the NCG immunodeficient mice model, clinical patients are assumed to have endogenous B cells and can also produce progenitor B cells over the course of their treatment. This production occurs in the bone marrow and consists of production of progenitor B cells that mature into adult B cells before trafficking into other tissues [**71**]. Another key difference between the patient model and the mouse model consists of implementing lymphodepletion. To account for lymphodepletion dynamics and the resulting longer-term B cell recovery, we incorporate a model concurrently developed by Parmar et al. that modulates both CAR-T and B-cell evolution using two parameters that evolve over the course of the lymphodepletion regimen: $DR_{dep}$ and $DR_{exp}$. The former decreases when there are fewer B cells in the system, and not only increases the rate of apoptosis in all relevant tissues but also reduces CAR-T metabolism from the liver compartment in

the early stages of B cell depletion. This slows down CAR-T depletion in the presence of B cells, important for that key early expansion. The latter parameter DRexp increases when there are fewer B cells and enhances the rate of nonspecific CAR-T expansion in the peripheral blood. Taken together, these model equations connects CAR-T proliferation and depletion with the patients' B-cell levels following the lymphodepletion regime in a manner consistent with experimental and clinical observations. All implementations are described in more detail in the methods section, and relevant equations are explicitly listed in Model Equations.

The resulting model, which includes adjusted tissue parameters, allometrically scaled tissue-specific transmigration rate constants, B cell dynamics, and lymphodepletion reasonably captures major characteristics of both CAR-T cellular kinetics and B cell depletion and recovery. Future work may further analyze individual covariates (e.g. age, initial tumor burden, etc.) and their precise effects on these blood profiles.

## 7.3. PBPK-PD model implementation

### 7.3.1. CART biodistributon and cellular kinetics

The main structure of the model comprises CAR-T biodistribution through various tissues ($n$ = 19). Similar to that of other published works [**65, 73, 67, 68**], most tissues in our model include both vascular and extravascular compartments through which CAR-T cells distribute. In the vascular compartment, CAR-T concentration over time is incremented by incoming blood flow, decremented by outgoing blood flow, and decremented by transmigration to the extravascular compartment as lymph (Figure 1a). In the corresponding extravascular compartment, CAR-T content is incremented by transmigration from the

vascular compartment and decremented by outgoing lymph flow.

$$V^{Vasc}\frac{dC^{Vasc}}{dt} = QC^{Vasc}_{Lung} - JC^{Vasc}V^{Vasc} - (Q-L)C^{Vasc} \tag{7.1}$$

$$V^{EVasc}\frac{dC^{EVasc}}{dt} = JC^{Vasc}V^{Vasc} - LC^{EVasc} \tag{7.2}$$

A few tissue types do not subscribe to this general rule. Blood from stomach, spleen, duodenum, and colon flow to the liver before returning to main circulation. Lymph from duodenum and colon flow to mesenteric lymph nodes (MLN) before flowing to the main lymphatic nodes (LN). In accordance with literature findings that there is no evidence of lymph flow in bone marrow, we omit an extravascular compartment in bone marrow [**74**]. Both LN and MLN are inherently lymphatic systems and do not have vascular compartments.

One unexpected but essential characteristic of the data is non-specific CAR-T expansion, evidenced by a rise in CAR-T concentration in NCG mice with no tumorous B cells ([**50**], Figure 2b) and similarly reported by *in vitro* assays [**69**]. To account for this observed, non-target driven expansion, we include a doubling term in the blood compartment so that CAR-Ts can proliferate passively. We assume that CAR T-cells have a natural proliferation rate with doubling time DT and some maximum steady-state concentration.

$$V_{Blood}\frac{dC_{Blood}}{dt} = \frac{\ln 2}{DT}\left(C_{Blood}V_{Blood}\left(1 - \frac{C_{Blood}}{C^{Max}_{Blood}}\right)\right) + ... \tag{7.3}$$

We considered alternatively implementing non-specific proliferation by either omitting a maximum concentration or by including this doubling component in all tissues, but we found that in both of those cases, the model diverged beyond reasonable biological values.

We also considered using a linear expansion rate rather than an exponential power law, but found that this did not sufficiently capture the observed proliferation in the data. To account for CAR-T metabolism and clearance from the body, we also include a linear depletion rate constant in the liver compartment.

$$V^E_{Liver}\frac{dC^E_{Liver}}{dt} = -K_{depletion} + ... \tag{7.4}$$

To make our model match with the available data, we converted model concentration units of cells per mL to copies of gDNA per $\mu g$, derived using information from [**75**].

$$[\text{copies gDNA}/\mu g] = 2039.37 \times [\text{cells}/mL]^{0.8537}. \tag{7.5}$$

### 7.3.2. Parameter selection

In constructing both the preclinical and clinical structural models, tissue volumes and flow rates for blood and lymph were set to reference values from literature when available [**73, 65, 64**]. For body measurements for which a datum was reported in mice or humans but not the other, we used allometric scaling exponents of −0.25 and 1 for volumes and flow rates, respectively, to estimate the missing datum. Across these sources, to address tissue types which blood flow was reported but not lymph flow, we estimated $L_{tissue}$ as 1/5000th of the corresponding blood flow. We further took total lymph flow $L_{LN}$ to be equal to the sum of all lymph flows and assumed there is no extravascular component of bone marrow.

Transmigration rate constants, which dictates the rate at which CAR-Ts diffuse from vascular to extravascular components of tissue, are taken from [**68**] when available. Transmigration rate constants for all other tissues were estimated by the model.

| | Description (Units) | Value | |
|---|---|---|---|
| | | Preclinical | Clinical |
| $J_{Lung}$ | Transmigration rate in lungs (1/hour) | 4705.13 | 654.03 |
| $J_{Spleen}$ | Transmigration rate in spleen (1/hour) | 1799.76 | 250.17 |
| $J_{Liver}$ | Transmigration rate in liver (1/hour) | 716.49 | 99.59 |
| $J_{Kidney}$ | Transmigration rate in kidney (1/hour) | 296.75 | 41.25 |

Table 7.1. First-order transmigration rate constants.

In implementing the B-cell distribution model, physiological parameters were taken from [**71**] and rate constants were estimated by [**70**].

| | Description (Units) | Value | |
|---|---|---|---|
| | | Preclinical | Clinical |
| $k_{B,prolif}$ | B cell proliferation rate (1/day) | N/A | .7 |
| $k_{B,apop}$ | B cell apoptotic rate (1/day) | N/A | .02 |
| $k_{B,mat}$ | Maturation rate of progenitor B cells (1/day) | N/A | .005 |
| $V_B$ | Volume of a B cell (mL) | $2 \times 10^{-10}$ | |
| $pb_{BM,ref}$ | Progenitor B cells concentration (cells/mL) | N/A | $7.9 \times 10^7$ |
| $B_{BM,ref}$ | Bone marrow B cell concentration (cells/mL) | $9.13 \times 10^{11}$ | $3.32 \times 10^8$ |
| $B_{PB,ref}$ | Peripheral blood B cell concentration (cells/mL) | $2.97 \times 10^9$ | $8.96 \times 10^5$ |
| $B_{SP,ref}$ | Spleen B cell concentration (cells/mL) | $7.01 \times 10^{11}$ | $4.02 \times 10^8$ |
| $B_{LN,ref}$ | Lymph node B cell concentration (cells/mL) | $3.76 \times 10^{11}$ | $1.55 \times 10^8$ |

Table 7.2. Parameters relevant for construction of B cell model. All rates are first-order rate constants. All concentrations are baseline values.

Since the mouse cohorts all comprise NCG immunodeficient mice, we assume they have no funtional immune system and as such, no progenitor B cell population. By similar logic, lymphodepletion only applies to clinical populations. Parameters related to lymphodepletion were estimated in a model by [**76**] and are used here.

| | Description (Units) | |
|---|---|---|
| | | Clinical |
| $KDE$ | Decay parameter (dimensionless) | .05 |
| $K_e$ | B cell regeneration rate constant (1/day) | 2.69 |
| $K_{out}$ | B cell depletion rate constant (1/day) | .0012 |
| $\gamma_{LD}$ | Lymphodepletion exponential factor | 1 |

Table 7.3. Parameters Associated with Lymphodepletion.

Lastly, in implementing dynamics dictating CAR-T target engagement, we use the following parameters from [**68**].

| | Description (Units) | Value | |
|---|---|---|---|
| | | Preclinical | Clinical |
| $\mathrm{DT}_{\mathrm{target}}$ | Nonspecific NALM-6 cell doubling time | 696 (hours) | 29 (days) |
| $\mathrm{Density}_{\mathrm{CAR}}$ | Density of CARs on CAR-T cells (number/cell) | 5000 | |
| $\mathrm{Density}_{\mathrm{Tumor}}$ | Density of TAA on tumor cells (number/cell) | 12590 | |
| $\mathrm{K}_{\mathrm{on}}$ | Binding affinity of CAR to TAA (1/Mol s) | $2.07 \times 10^5$ | |
| $\mathrm{K}_{\mathrm{off}}$ | Dissociation rate of CAR to TAA (1/day) | $6.81 \times 10^{-5}$ | |

Table 7.4. Parameters associated with CAR-Target Engagement.

### 7.3.3. Characterizing CAR-T biodistribution with no target engagement

We first developed a PBPK-PD model to characterize the biodistribution of CAR-T cells in a control group of NCG immunodeficient mouse strain with no intravenously-injected B cells. Estimated parameters include non-specific doubling time, maximum CAR-T concentration in the blood, and first-order transmigration rate constants. Notably, this first iteration of the model included no CAR-T target engagement, only passive proliferation in the blood and linear depletion through the liver. Fitting this model to observed data yields reasonably accurate model fitted profiles (Figure 7.2) and estimates of unknown transmigration rate constants.

### 7.3.4. Incorporating target-engagement into the existing model

With these estimated rate constants from fitting the initial model, we then expand the model to account for data from an NCG immunodeficient mouse strain with $10^6$ NALM-6 cells injected intravenously. Since these mice have no functional immune system, we assume that there is no innate production of B cells within bone marrow, and all B cells in the system are those that were injected.

### 7.3.5. Preclinical to clinical scaling

Since the dataset does not include tissue-specific data for clinical patients, we cannot use the model to estimate parameters for transmigration rate constants. To achieve those values, we scale the preclinical estimates with an allometric exponent of −0.25.

$$J_{human} = J_{mouse} \times \left(\frac{\text{human mass}}{\text{mouse mass}}\right)^{-0.25} \tag{7.6}$$

All other parameters, including tissue volumes and flow rates, are taken from literature.

### 7.3.6. Model-estimated values

The main purpose of model simulation was to estimate parameters relevant for CAR-T expansion and killing efficacy, which together can give mechanistic insights towards how their mechanisms scale across species. The expansion rate is governed by two parameters: $K_{exp,max}$ and $EC50_{exp}$, which dictate the maximum expansion rate and sensitivity of expansion (Hill curve characteristics), respectively. The killing rate is similarly governed by parameters $K_{kill,max}$ and $KC50_{kill}$, which dictate the maximum killing rate and sensitivity to effector concentration, respectively.

| | Description (Units) | Preclinical (Error %) | Clinical (Error %) |
|---|---|---|---|
| $K_{exp,max}$ | Maximum expansion rate constant (1/day) | .917 (6.38) | 4.33 (6.81) |
| $EC50_{exp}$ | Complex/CART ratio achieving half $K_{exp,max}$ | 1.15 (2.13) | 5.09 (102) |
| $K_{kill,max}$ | Maximum killing rate constant (1/day) | .026 (24.9) | .55 (24.9) |
| $KC50_{kill}$ | Complex Tumor ratio achieving half $K_{kill,max}$ | 2.25 (102) | $3 \times 10^{-6}$ (1e3) |

Table 7.5. Model-estimated parameters for CAR-Target engagement

Estimated values suggest key mechanistic differences between preclinical to clinical pharmacology, such as much higher expansion rates and killing rates in clinical patients compared to mice, although further fine-tuning can refine this specific model and increase its utility in advising clinical dosing strategies. This modeling effort also provides a framework for future PBPK-PD models that seek to address preclinical-to-clinical translation of various treatments and drug modalities. Future efforts may also validate these results against known translational differences or other CAR-T products.

## 7.4. Conclusion and discussion

This work establishes a modeling framework that can be used to explore translation of CAR-T cell therapies. The model simultaneously fits data from a cohort of non-tumor-bearing mice, tumor-bearing mice, and clinical cancer patients, and in doing so estimates biophysical parameters relevant for killing and expansion across each of the populations. Although there is existing work addressing preclinical-to-clinical translation of CAR-T cells between mice and patients, these models focus on BCMA and EGFR-targeting products and use substantially fewer anatomical compartments, which limits

their precision and sensitivity towards individual variability.. These studies also do not differentiate between non-tumor-bearing and tumor-bearing mouse models, from which the datasets provide insight on nonspecific expansion.

The model accounts the interactions between dynamic populations of CAR-T cells and CD19-expressing B lymphocytes. They each have individual growth rates and interact to form CAR-Target complexes; these complexes mediate the killing of tumor cells which enables CAR-Ts to expand. For simplicity, we omit CAR-T exhaustion from the model, but future work may address exhaustion, a biological phenomenon especially pertinent to high-affinity CAR-T cells [**77**]. Future avenues of expansion (no pun intended) include incorporating different functional roles of effector and memory T-cells, or incorporating mechanisms of cytokine production. It would also be straightforward to extend the existing analyses to account for other attributes of CAR-T cell therapy – CAR affinity (which can dictate how quickly CAR-T cells exhaust), construct-specific modalities, cancer types beyond hematological malignancies, cancer severity and other covariates, and lymphodepletion dosing strategies – or to other biologics, such as CAR-NKs that have demonstrated similar clinical success [**78**].

This work used an experimental dataset on CD19-targeting CAR-T cells as a case study and expanded a previous modeling framework to include more highly detailed anatomy more biophysically accurate B cell time evolution dynamics, and key lymphodepletion characteristics. It also addresses differences interspecies differences in expansion and cell-killing efficacy for this CAR-T product. Taken together, this effort furthers our collective understanding of how dose-exposure-response relationships scale between

species. It simultaneously captures dynamics of CAR-T biodistribution and CK in immunocompromised NCG mice (both with-and without disseminated NALM-6 cells) and CK and lymphodepletion in clinical patients. The suggested mechanistic differences can inform clinical dosing strategies for existing products, as well as predict clinical PKPD characteristics for drugs that have yet been tested in the clinic. The modeling framework itself can also be used to study preclinical-to-clinical translation of other CAR-T products or cell therapies.

## 7.5. Model equations

The model describes biodistribution and cellular kinetics of CAR-T cells in mice and cancer patients. It includes species-specific tissue volumes, flow rates (for peripheral blood and lymphatic tissues), differential kinetics of CAR-T cells, allometrically-scaled transmigration rates, non-specific and target-mediated expansion of CAR-T cells and target-engagement driven B-cell depletion, and linear depletion through the liver.

In preclinical mice experiments, both CAR-T cells and target NALM-6 tumor cells are injected intravenously through the tail vein. To that end, there is no tumor compartment as in previous models [**73**], and B cells diffuse between the peripheral blood, spleen, liver, and bone marrow. We assume that the preclinical mouse lines (immunodeficient NCG mice) have neither endogenous B cells nor progenitor B cells. To address anatomical differences in mouse sex, we include a boolean variable within the dataset that denotes whether each datum comes from male mice (for which there is data on testes and epididymis biodistribution) or female (ovaries and uteri).

The clinical patients comprise a mix of DLBCL, FL, and MZL patients, and we represent their initial tumor burden using baseline cancerous B cell concentrations from [**71**]. Both preclinical and clinical model equations suppose that B-cell-engagement CAR-T expansion occurs in a few select tissues, whereas nonspecific CAR-T expansion occurs in the peripheral blood and flows into the other tissues.

### 7.5.1. Cellular kinetics and biodistribution equations

#### Peripheral Blood

(7.7)

$$
\begin{aligned}
V_{\text{Blood}}\frac{dC_{\text{Blood}}}{dt} = &-\left(Q_{\text{Lung}} - L_{\text{Lung}}\right)C_{\text{Whole Blood}} + L_{\text{LN}}C_{\text{LN}} \\
&+ \left(Q_{\text{Spleen}} - L_{\text{Spleen}} + Q_{\text{Stomach}} - L_{\text{Stomach}} + Q_{\text{Duodenum}} - L_{\text{Duodenum}}\right)C_{\text{Liver}}^{\text{Vasc}} \\
&+ \left(Q_{\text{Colon}} - L_{\text{Colon}} + Q_{\text{Liver}} - L_{\text{Liver}}\right)C_{\text{Liver}}^{\text{Vasc}} \\
&+ \left(Q_{\text{Kidney}} - L_{\text{Kidney}}\right)C_{\text{Kidney}}^{\text{Vasc}} + \left(Q_{\text{Brain}} - L_{\text{Brain}}\right)C_{\text{Brain}}^{\text{Vasc}} \\
&+ \left(Q_{\text{Heart}} - L_{\text{Heart}}\right)C_{\text{Heart}}^{\text{Vasc}} + \left(Q_{\text{Muscle}} - L_{\text{Muscle}}\right)C_{\text{Muscle}}^{\text{Vasc}} \\
&+ \left(Q_{\text{Adipose}} - L_{\text{Adipose}}\right)C_{\text{Adipose}}^{\text{Vasc}} + \left(Q_{\text{Bone marrow}} - L_{\text{Bone marrow}}\right)C_{\text{Bone marrow}}^{\text{Vasc}} \\
&+ \left(Q_{\text{Uterus/Testes}} - L_{\text{Uterus/Testes}}\right)C_{\text{Uterus/Testes}}^{\text{Vasc}} \\
&+ \left(Q_{\text{Ovary/Epididymis}} - L_{\text{Ovary/Epididymis}}\right)C_{\text{Ovary/Epididymis}}^{\text{Vasc}} \\
&+ \left(Q_{\text{Other}} - L_{\text{Other}}\right)C_{\text{Other}}^{\text{Vasc}} \\
&+ \text{Hillexp}_{\text{Blood}}C_{\text{Blood}} + DR_{exp}C_{\text{Blood}}
\end{aligned}
$$

**Lung**

$$V_{\text{Lung}}^{\text{Vasc}} \frac{dC_{\text{Lung}}^{\text{Vasc}}}{dt} = \left(Q_{\text{Lung}} - L_{\text{Lung}}\right) C_{\text{Whole Blood}} - J_{\text{Lung}} C_{\text{Lung}}^{\text{Vasc}} V_{\text{Lung}}^{\text{Vasc}} - Q_{\text{Lung}} C_{\text{Lung}}^{\text{Vasc}} \tag{7.8}$$

$$V_{\text{Lung}}^{\text{EVasc}} \frac{dC_{\text{Lung}}^{\text{EVasc}}}{dt} = J_{\text{Lung}} C_{\text{Lung}}^{\text{Vasc}} V_{\text{Lung}}^{\text{Vasc}} - L_{\text{Lung}} C_{\text{Lung}}^{\text{EVasc}} \tag{7.9}$$

**Liver**

$$\begin{aligned} V_{\text{Liver}}^{\text{Vasc}} \frac{dC_{\text{Liver}}^{\text{Vasc}}}{dt} &= Q_{\text{Liver}} C_{\text{Lung}}^{\text{Vasc}} - J_{\text{Liver}} C_{\text{Liver}}^{\text{Vasc}} V_{\text{Liver}}^{\text{Vasc}} \\ &+ \left(Q_{\text{Spleen}} - L_{\text{Spleen}}\right) C_{\text{Spleen}}^{\text{Vasc}} + \left(Q_{\text{Stomach}} - L_{\text{Stomach}}\right) C_{\text{Stomach}}^{\text{Vasc}} \\ &+ \left(Q_{\text{Duodenum}} - L_{\text{Duodenum}}\right) C_{\text{Duodenum}}^{\text{Vasc}} + \left(Q_{\text{Colon}} - L_{\text{Colon}}\right) C_{\text{Colon}}^{\text{Vasc}} \\ &- \left(Q_{\text{Spleen}} - L_{\text{Spleen}} + Q_{\text{Stomach}} - L_{\text{Stomach}}\right) C_{\text{Liver}}^{\text{Vasc}} \\ &- \left(Q_{\text{Duodenum}} - L_{\text{Duodenum}} + Q_{\text{Colon}} - L_{\text{Colon}}\right) C_{\text{Liver}}^{\text{Vasc}} \\ &- \left(Q_{\text{Liver}} - L_{\text{Liver}}\right) C_{\text{Liver}}^{\text{Vasc}} \end{aligned} \tag{7.10}$$

$$\begin{aligned} \frac{dC_{\text{Liver}}^{\text{EVasc}}}{dt} &= \left(J_{\text{Liver}} C_{\text{Liver}}^{\text{Vasc}} V_{\text{Liver}}^{\text{Vasc}} - L_{\text{Liver}} C_{\text{Liver}}^{\text{EVasc}} - K_{\text{Liver}} C_{\text{Liver}}^{\text{EVasc}} V_{\text{Liver}}^{\text{EVasc}}\right) / V_{\text{Liver}}^{\text{EVasc}} \\ &- K_{dep,\text{Liver}} C_{\text{Liver}}^{\text{EVasc}} - K_{dep\text{LD}} DR_{dep} C_{\text{Liver}}^{\text{EVasc}} \end{aligned} \tag{7.11}$$

**Lymph Node**

(7.12)

$$\begin{aligned} V_{\mathrm{LN}}\frac{dC_{\mathrm{LN}}}{dt} &= L_{\mathrm{Lung}}C^{\mathrm{EVasc}}_{\mathrm{Lung}} + L_{\mathrm{Liver}}C^{\mathrm{EVasc}}_{\mathrm{Liver}} + L_{\mathrm{Heart}}C^{\mathrm{EVasc}}_{\mathrm{Heart}} + L_{\mathrm{Muscle}}C^{\mathrm{EVasc}}_{\mathrm{Muscle}} + L_{\mathrm{Skin}}C^{\mathrm{EVasc}}_{\mathrm{Skin}} \\ &+ L_{\mathrm{Adipose}}C^{\mathrm{EVasc}}_{\mathrm{Adipose}} + L_{\mathrm{Bone\ marrow}}C^{\mathrm{EVasc}}_{\mathrm{Bone\ marrow}} + L_{\mathrm{Brain}}C^{\mathrm{EVasc}}_{\mathrm{Brain}} + L_{\mathrm{Kidney}}C^{\mathrm{EVasc}}_{\mathrm{Kidney}} \\ &+ L_{\mathrm{Spleen}}C^{\mathrm{EVasc}}_{\mathrm{Spleen}} + L_{\mathrm{Stomach}}C^{\mathrm{EVasc}}_{\mathrm{Stomach}} + L_{\mathrm{Other}}C^{\mathrm{EVasc}}_{\mathrm{Other}} \\ &+ L_{\mathrm{Tumor}}\left(C^{\mathrm{EVasc}}_{\mathrm{Tumor}} - \frac{C_{\mathrm{TE}}}{Density_{\mathrm{CAR}}}\right) \\ &+ L_{\mathrm{MLN}}C_{\mathrm{MLN}} - L_{\mathrm{LN}}C_{\mathrm{LN}} \\ &+ \mathrm{Hillexp}_{\mathrm{Blood}}C_{\mathrm{Blood}}V_{\mathrm{LN}} \end{aligned}$$

**Mesenteric Lymph Node**

$$V_{\mathrm{MLN}}\frac{dC_{\mathrm{MLN}}}{dt} = L_{\mathrm{Duodenum}}C^{\mathrm{EVasc}}_{\mathrm{Duodenum}} + L_{\mathrm{Colon}}C^{\mathrm{EVasc}}_{\mathrm{Colon}} - L_{\mathrm{MLN}}C_{\mathrm{MLN}} \tag{7.13}$$

**Other Tissues**

$$V^{\mathrm{Vasc}}_{\mathrm{Tissue}}\frac{dC^{\mathrm{Vasc}}_{\mathrm{Tissue}}}{dt} = Q_{\mathrm{Tissue}}C^{\mathrm{Vasc}}_{\mathrm{Lung}} - J_{\mathrm{Tissue}}C^{\mathrm{Vasc}}_{\mathrm{Tissue}}V^{\mathrm{Vasc}}_{\mathrm{Tissue}} - \left(Q_{\mathrm{Tissue}} - L_{\mathrm{Tissue}}\right)C^{\mathrm{Vasc}}_{\mathrm{Tissue}} \tag{7.14}$$

$$V^{\mathrm{EVasc}}_{\mathrm{Tissue}}\frac{dC^{\mathrm{EVasc}}_{\mathrm{Tissue}}}{dt} = J_{\mathrm{Tissue}}C^{\mathrm{Vasc}}_{\mathrm{Tissue}}V^{\mathrm{Vasc}}_{\mathrm{Tissue}} - L_{\mathrm{Spleen}}C^{\mathrm{EVasc}}_{\mathrm{Tissue}} \tag{7.15}$$

### 7.5.2. Lymphodepletion

Each of the 13 cancer patients underwent a lymphodepletion regimen of 25 $mg/m^2$ of fludarabine and 250 $mg/m^2$ of cyclophosphamide. We represent lymphodepletion and the ensuing B-cell recovery with the following parameters and constrained the model to B cell

aplasia clinical data. This part of the model was concurrently developed by [**76**].

Lymphodepletion equation parameters:

$$KDE = 0.05 \tag{7.16}$$

$$K_e = 2.69 \tag{7.17}$$

$$K_{\text{out}} = .0012 \tag{7.18}$$

$$K_{\text{in}} = K_{\text{out}} \times B_{\text{PB,ref}} \tag{7.19}$$

$$\gamma_{LD} = 1 \tag{7.20}$$

Lymphodepletion initial conditions and equations.

$$A_0 = 0 \tag{7.21}$$

$$R_0 = B_{\text{PB,ref}} \tag{7.22}$$

$$DR_{\text{exp}} = \frac{B_{\text{PB,ref}} - R}{B_{\text{PB,ref}}} \tag{7.23}$$

$$DR_{\text{dep}} = \left(\frac{R}{B_{\text{PB,ref}}}\right)^{\gamma_{LD}} \tag{7.24}$$

$$\frac{dA}{dt} = -KDE \times A \tag{7.25}$$

$$\frac{dR}{dt} = K_{\text{in}} - K_{\text{out}} R\left(1 + K_e A\right) \tag{7.26}$$

### 7.5.3. B cell dynamics and CAR-target complexes

These parameters and equations were adapted from the B cell model concurrently developed by Kesioglou and Maddah [**70**].

Rate constants:

$$k_B\text{exit} = .465, \qquad k_B\text{prolif} = 7, \qquad k_B\text{apop} = .02, \qquad k_B\text{mat} = .25 \times k_B\text{apop}$$

Reference B cell concentrations and partition coefficients: The first term $pb_{\text{BM,ref}}$ is the reference concentration of progenitor B cells in the bone marrow and is assumed to be zero in the preclinical mouse lines. Initial conditions for each B cell concentration

$$pb_{\text{BM,ref}} = 5 \times 10^9 \text{cells}/V_{\text{Bonemarrow}} \tag{7.27}$$

$$B_{\text{BM,ref}} = 2.1 \times 10^{10} \text{cells}/V_{\text{Bonemarrow}} \tag{7.28}$$

$$B_{\text{PB,ref}} = 2.8 \times 10^9 \text{cells}/V_{\text{Blood}} \tag{7.29}$$

$$B_{\text{SP,ref}} = 2.86 \times 10^{10} \text{cells}/\left(V_{\text{Spleen}}^{\text{Vasc}} + V_{\text{Spleen}}^{\text{EVasc}}\right) \tag{7.30}$$

$$B_{\text{LN,ref}} = 4.25 \times 10^{10} \text{cells}/V_{\text{LN}} \tag{7.31}$$

$$K_{\text{BM,B,ref}} = B_{\text{BM,ref}}/B_{\text{PB,ref}} \tag{7.32}$$

$$K_{\text{SP,B,ref}} = B_{\text{SP,ref}}/B_{\text{PB,ref}} \tag{7.33}$$

$$K_{\text{LN,B,ref}} = B_{\text{LN,ref}}/B_{\text{PB,ref}} \tag{7.34}$$

Generation, apoptosis, proliferation, and maturation rates in the bone marrow:

$$\text{Gen\_rate} = pb_{\text{BM}} \times (k_B\text{apop} + k_B\text{mat}) \times \max\left(0, 1 - \frac{B_{\text{BM}}}{B_{\text{BM,ref}}}\right) \tag{7.35}$$

$$\text{Apop\_rate} = k_B\text{apop} \times pb_{BM} \tag{7.36}$$

$$\text{Prolif\_rate} = pb_{\text{BM}} \times k_B\text{prolif} \times \max\left(0, 1 - \frac{pb_{\text{BM}}}{pb_{\text{BM,ref}}}\right) \tag{7.37}$$

$$\text{Mat\_rate} = k_B\text{mat} \times pb_{\text{BM}} \tag{7.38}$$

$$\frac{d\left(pb_{\text{BM}}\right)}{dt} = \text{Gen\_rate} - \text{Apop\_rate} + \text{Prolif\_rate} - \text{Mat\_rate} \tag{7.39}$$

B cell trafficking within and between tissues:

$$\text{Apop\_tissue} = k_B\text{apop} \times B_{\text{tissue}} \times DR_{\text{dep}} \tag{7.40}$$

$$\text{Prolif\_tissue} = k_{\text{B}}\text{prolif} \times \max\left(0, 1 - \frac{B_{\text{tissue}}}{B_{\text{tissue,ref}}}\right) \tag{7.41}$$

$$\text{Traffic}_{\text{BM,PB}} = k_B\text{exit} \times \left(\frac{B_{\text{BM}}}{K_{\text{BM,B,ref}}} - B_{\text{PB}}\right) \tag{7.42}$$

$$\text{Traffic}_{\text{PB,SP}} = k_B\text{exit} \times \left(B_{\text{PB}} - \frac{B_{\text{SP}}}{K_{\text{SP,B,ref}}}\right) \tag{7.43}$$

$$\text{Traffic}_{\text{PB,LN}} = k_B\text{exit} \times \left(B_{\text{PB}} - \frac{B_{\text{LN}}}{K_{\text{LN,B,ref}}}\right) \tag{7.44}$$

$$\begin{aligned}\frac{d\left(B_{\text{blood}}\right)}{dt} &= \text{Mat\_rate} - \text{Apop\_rate} + \text{Prolif\_rate} - \text{Traffic} \\ &\quad - \text{Hillkill}_{\text{tissue}} \times B_{\text{blood}} - K_{\text{out}}\left(1 - K_e A\right) B_{\text{blood}}\end{aligned} \tag{7.45}$$

$$\begin{aligned}\frac{d\left(B_{\text{tissue}}\right)}{dt} &= \text{Mat\_rate} - \text{Apop\_rate} + \text{Prolif\_rate} \\ &\quad - \text{Traffic} - \text{Hillkill}_{\text{tissue}} \times B_{\text{tissue}}\end{aligned} \tag{7.46}$$

### 7.5.4. Target-mediated expansion and killing:

$C_{\text{tissue}}$ refers to CAR-T concentration in the tissue, $B_{\text{tissue}}$ refers to B cell concentration in the tissue, and $Cplx_{\text{tissue}}$ refers to CAR-target complex concentration in the tissue.

$$\text{Hillexp}_{\text{tissue}} = \frac{K_{\text{exp,max}} \times \frac{\text{Cplx}_{\text{tissue}}}{C_{\text{Blood}}}}{\frac{\text{Cplx}_{\text{tissue}}}{C_{\text{tissue}}} + EC_{50}} \tag{7.47}$$

$$\text{Hillkill}_{\text{tissue}} = \frac{K_{\text{kill,max}} \times \frac{\text{Cplx}_{\text{tissue}}}{B_{\text{tissue}}}}{\frac{\text{Cplx}_{\text{tissue}}}{\text{B}_{\text{tissue}}} + KC_{50}} \tag{7.48}$$

$$\begin{aligned}\frac{d}{dt}\text{Cplx}_{\text{tissue}} = {} & K_{on}\left(\text{Density}_{\text{CAR}}C_{\text{tissue}} - \text{Cplx}_{\text{tissue}}\right)\left(\text{Density}_{\text{TAA}}B_{\text{tissue}} - \text{Cplx}_{\text{tissue}}\right) \\ & - K_{off}\text{Cplx}_{\text{tissue}}\end{aligned} \tag{7.49}$$

**Modeling Software**: All model fittings were performed using the Stochastic Approximation Expectation Maximization (SAEM) algorithm within Monolix 2020R1 (Lixoft ®)

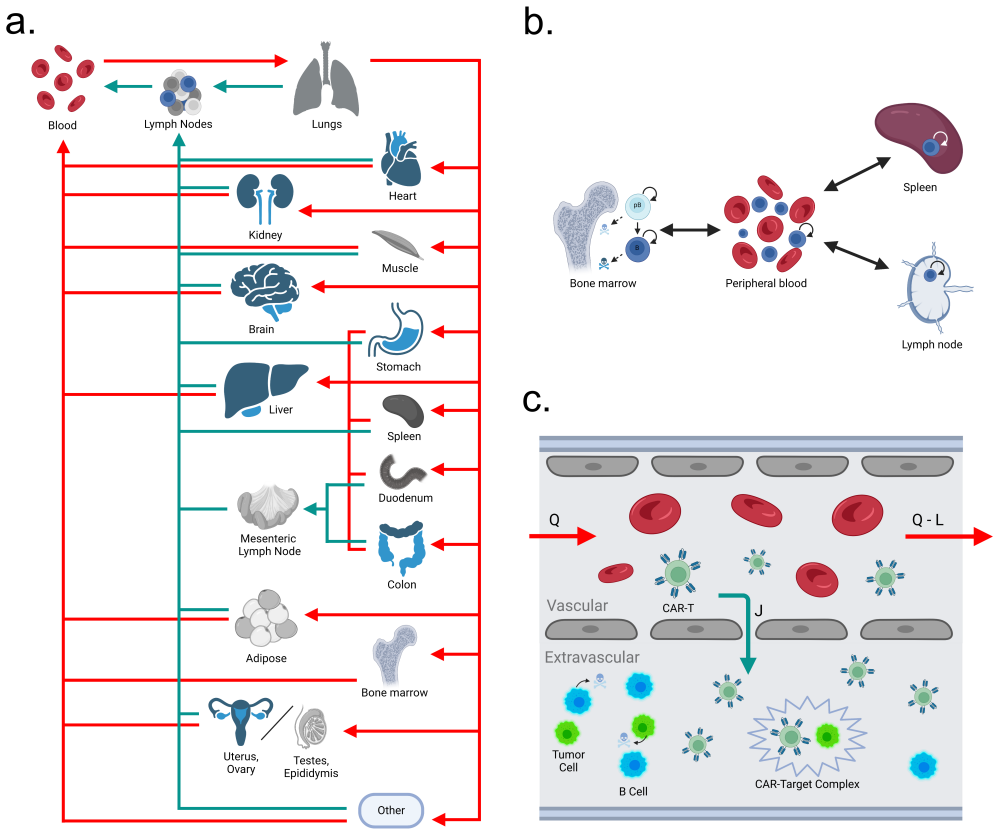


Figure 7.1. (a) Blood and lymph flows through pertinent organs ($n$ = 20). Solid red arrows indicate blood flow, and teal arrows indicate lymph flow. For most tissues, a portion of incoming blood travels through the extravascular compartment as lymph, flowing through the lymph nodes before returning to the main blood pool. Similarly, lymph from duodenum and colon flow through the mesenteric lymph node (MLN) before returning to the main lymph node compartment. The "Other" compartment comprises issues for which we do not have data. (b) B cell production and distribution. Progenitor B cells (light blue circles) are produced in the bone marrow and mature into full B cells (dark blue circles), which distribute between the peripheral blood, lymph node, and spleen. All flow rates are bidirectional. (c) CAR-T movement and target engagement in the vascular and extravascular compartment of sample tissue. CAR-T cells form complexes with target tumor cells (bright green).

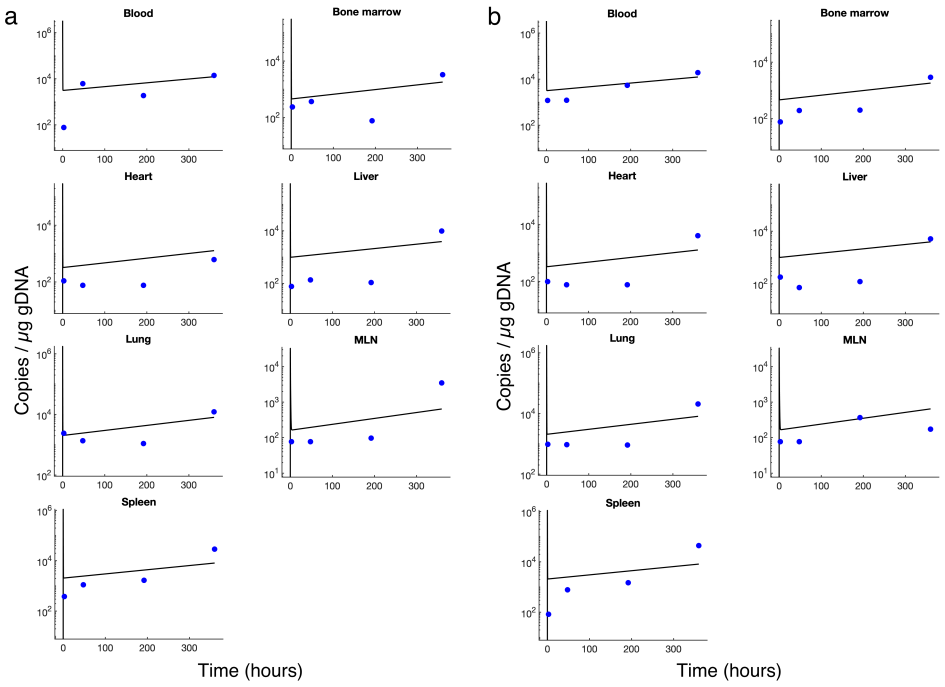


Figure 7.2. Model fitted profiles of CAR-T biodistribution in non-tumor-bearing NCG (a) male and (b) female mice. Solid blue dots indicate observed data and black lines indicate model fits. Data show CAR-T concentration (in copies per $\mu g$ of genomic DNA) versus hours after intravenous infusion of $5 \times 10^6$ CAR-T cells. Observed data was digitized from the original study and show averaged values in (a) male and (b) female mice. Each set of data is the result of averaging over three mice.

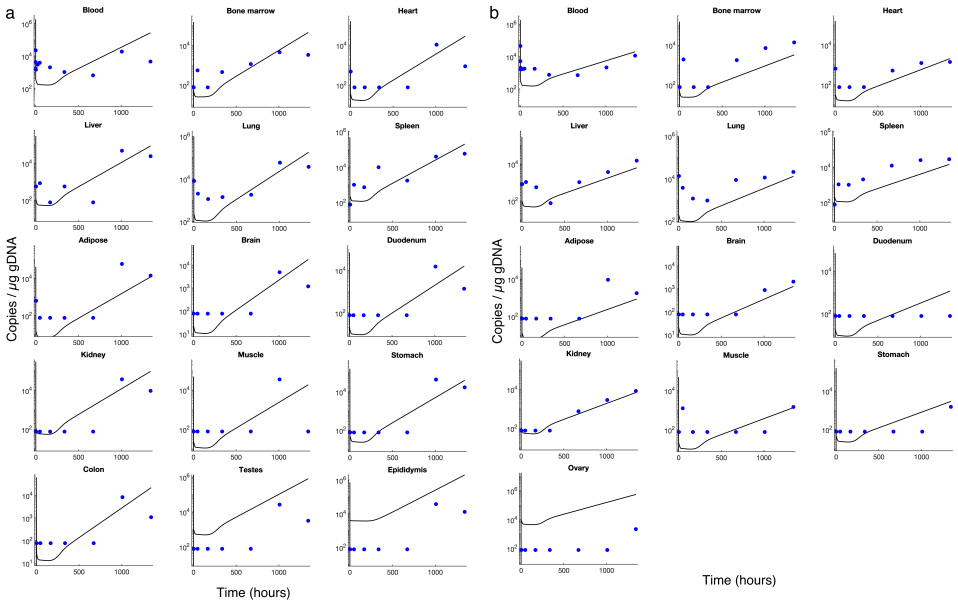


Figure 7.3. Model fitted profiles of CAR-T biodistribution in tumor-bearing NCG (a) male and (b) female mice. Solid blue dots indicate observed data and black lines indicate model fits. Data show CAR-T concentration (in copies per $\mu g$ of genomic DNA) versus hours after intravenous infusion of $5 \times 10^6$ CAR-T cells. Observed data was digitized from the original study and show averaged values over three mice.

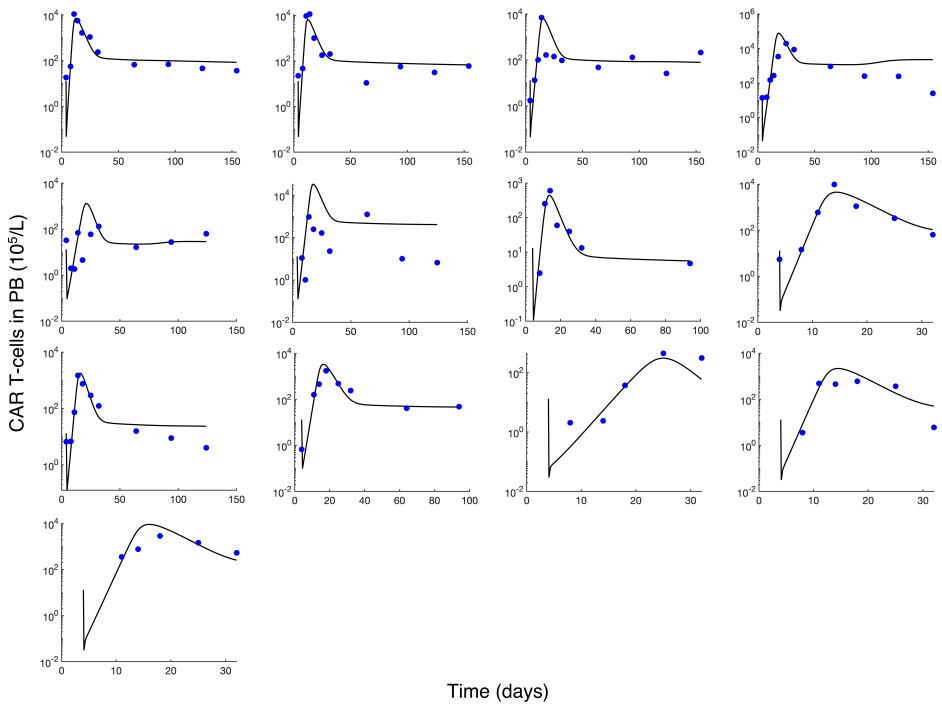


Figure 7.4. Model fitted profiles of CAR-T cellular kinetics in 13 clinical patients. Solid blue dots indicate observed data and black lines indicate model fits. Figure shows concentration, measured in CAR-T cell concentration in peripheral blood (in $10^5$ cells/L), versus days post CAR-T infusion. All patients received a intravenous dose of $10^6$ CAR-T cells per kg of body weight.

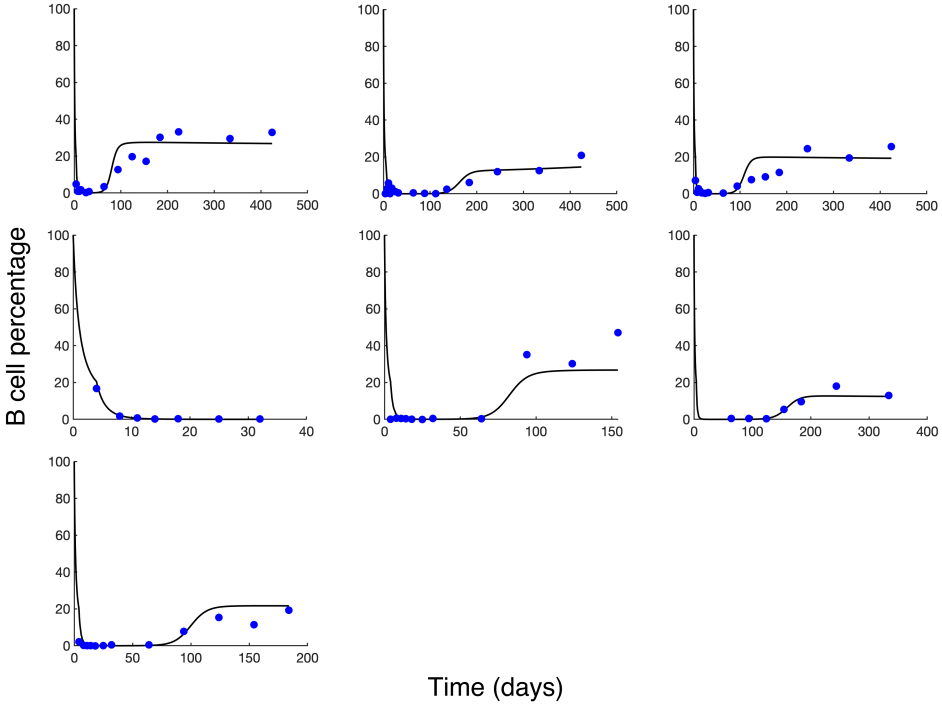


Figure 7.5. Model fitted profiles of B-cell recovery in 13 clinical patients. Solid blue dots indicate observed data and black lines indicate model fits. Figure shows concentration, measured in B-cell percentage of baseline, versus days post CAR-T infusion.

# References

[1] C. Bushdid, M. O. Magnasco, L. B. Vosshall, and A. Keller. Humans can discriminate more than 1 trillion olfactory stimuli. *Science*, 343:1370–1372, 2014.

[2] Richard C Gerkin and Jason B Castro. The number of olfactory stimuli that humans can discriminate is still unknown. *eLife*, 2015.

[3] Markus Meister. On the dimensionality of odor space. *eLife*, 2015.

[4] Birte L. Nielsen, Tadeusz Jezierski, J. Elizabeth Bolhuis, Luisa Amo, Frank Rosell, Marije Oostindjer, Janne W. Christensen, Dorothy McKeegan, Deborah L. Wells, and Peter Hepper. Olfaction: An overlooked sensory modality in applied ethology and animal welfare. *Frontiers in Veterinary Science*, 2, 12 2015.

[5] Shin Nagayama, Ryota Homma, and Fumiaki Imamura. Neuronal organization of olfactory bulb circuits. *Frontiers in Neural Circuits*, 8, 9 2014.

[6] Monica W. Chu, Wankun L. Li, and Takaki Komiyama. Balancing the robustness and efficiency of odor representations during learning. *Neuron*, 92:174–186, 10 2016.

[7] Olivier Gschwend, Nixon M. Abraham, Samuel Lagier, Frédéric Begnaud, Ivan Rodriguez, and Alan Carleton. Neuronal pattern separation in the olfactory bulb improves odor discrimination learning. *Nature Neuroscience*, 18:1474–1482, 10 2015.

[8] Maxim Bazhenov and Mark Stopfer. Forward and back: Motifs of inhibition in olfactory processing. *Neuron*, 67:357–358, 8 2010.

[9] Anne Grelat, Laura Benoit, Sébastien Wagner, Carine Moigneu, Pierre-Marie Lledo, and Mariana Alonso. Adult-born neurons boost odor-reward association. *Proceedings of the National Academy of Sciences*, 115:2514–2519, 2018.

[10] Wankun L Li, Monica W Chu, An Wu, Yusuke Suzuki, Itaru Imayoshi, and Takaki Komiyama. Adult-born neurons facilitate olfactory bulb pattern separation during task engagement. 7:33006, 2018.

[11] Daniel Nunes and Thomas Kuner. Disinhibition of olfactory bulb granule cells accelerates odour discrimination in mice. *Nature Communications*, 6, 11 2015.

[12] Ramon Bartolo, Richard C. Saunders, Andrew R. Mitz, and Bruno B. Averbeck. Information-limiting correlations in large neural populations. *Journal of Neuroscience*, 40:1668–1678, 2 2020.

[13] Chengcheng Huang, Alexandre Pouget, and Brent Doiron. Internally generated population activity in cortical networks hinders information transmission. *Science Advances*, 8, 2022.

[14] Mohammad Mehdi Kafashan, Anna W. Jaffe, Selmaan N. Chettih, Ramon Nogueira, Inigo Arandia-Romero, Christopher D. Harvey, Ruben Moreno-Bote, and Jan Drugowitsch. Scaling of sensory information in large neural populations shows signatures of information-limiting correlations. *Nature Communications*, 12, 12 2021.

[15] Rubén Moreno-Bote, Jeffrey Beck, Ingmar Kanitscheider, Xaq Pitkow, Peter Latham, and Alexandre Pouget. Information-limiting correlations. *Nature Neuroscience*, 17:1410–1417, 10 2014.

[16] Peggy Seriès, Peter E. Latham, and Alexandre Pouget. Tuning curve sharpening for orientation selectivity: Coding efficiency and the impact of correlations. *Nature Neuroscience*, 7:1129–1135, 10 2004.

[17] L F Abbott and Peter Dayan. The effect of correlated variability on the accuracy of a population code. *Neural Computation*, 11:91–101, 1999.

[18] Haim Sompolinsky, Hyoungsoo Yoon, Kukjin Kang, and Maoz Shamir. Population coding in neuronal systems with correlated noise. *Physical Review E*, 64:11, 2001.

[19] Volker Pernice and Rava Azeredo da Silveira. Interpretation of correlated neural variability from models of feed-forward and recurrent circuits. *PLoS Computational Biology*, 14, 2 2018.

[20] Martina Valente, Giuseppe Pica, Giulio Bondanelli, Monica Moroni, Caroline A. Runyan, Ari S. Morcos, Christopher D. Harvey, and Stefano Panzeri. Correlations enhance the behavioral readout of neural population activity in association cortex. *Nature Neuroscience*, 24:975–986, 7 2021.

[21] Robert Meyer, Josef Ladenbauer, and Klaus Obermayer. The influence of mexican hat recurrent connectivity on noise correlations and stimulus encoding. *Frontiers in Computational Neuroscience*, 11, 5 2017.

[22] Shawn D. Burton and Nathaniel N. Urban. Cell and circuit origins of fast network oscillations in the mammalian main olfactory bulb. *eLife*, 10, 10 2021.

[23] Sonya Giridhar, Brent Doiron, and Nathaniel N. Urban. Timescale-dependent shaping of correlation by olfactory bulb lateral inhibition. *Proceedings of the National Academy of Sciences*, 108:5843–5848, 4 2011.

[24] Mario Trevino. Inhibition controls asynchronous states of neuronal networks. *Frontiers in Synaptic Neuroscience*, 8, 2016.

[25] Cheng Ly, Andrea K. Barreiro, Shree Hari Gautam, and Woodrow L. Shew. Odor-evoked increases in olfactory bulb mitral cell spiking variability. *iScience*, 24, 9 2021.

[26] Nicolas Fourcaud-Trocmé, David Hansel, Carl Van Vreeswijk, and Nicolas Brunel. How spike generation mechanisms determine the neuronal response to fluctuating inputs. *The Journal of Neuroscience*, 23:11628–11640, 2003.

[27] Nixon M. Abraham, Hartwig Spors, Alan Carleton, Troy W. Margrie, Thomas Kuner, and Andreas T. Schaefer. Maintaining accuracy at the expense of speed: Stimulus similarity defines odor discrimination time in mice. *Neuron*, 44:865–876, 12 2004.

[28] Nixon M. Abraham, Delphine Guerin, Khaleel Bhaukaurally, and Alan Carleton. Similar odor discrimination behavior in head-restrained and freely moving mice. *PLoS ONE*, 7, 12 2012.

[29] Monica W. Chu, Wankun L. Li, and Takaki Komiyama. Lack of pattern separation in sensory inputs to the olfactory bulb during perceptual learning. *eNeuro*, 4, 9 2017.

[30] Felix Franke, Rodrigo Quian Quiroga, Andreas Hierlemann, and Klaus Obermayer. Bayes optimal template matching for spike sorting - combining fisher discriminant analysis with optimal filtering. *Journal of Computational Neuroscience*, 38:439–459, 6 2015.

[31] Christopher D. Wilson, Gabriela O. Serrano, Alexei A. Koulakov, and Dmitry Rinberg. A primacy code for odor identity. *Nature Communications*, 8, 12 2017.

[32] Claudia Clopath, Lars Busing, Eleni Vasilaki, and Wulfram Gerstner. Connectivity reflects coding: A model of voltage-based STDP with homeostasis. *Nature Neuroscience*, 13:344–352, 3 2010.

[33] Hongyu Meng and Hermann Riecke. Structural spine plasticity in olfaction: Memory and forgetting, enhanced vs. reduced discriminability after learning, 12 2020.

[34] Kurt A. Sailor, Matthew T. Valley, Martin T. Wiechert, Hermann Riecke, Gerald J. Sun, Wayne Adams, James C. Dennis, Shirin Sharafi, Guo li Ming, Hongjun Song, and Pierre Marie Lledo. Persistent structural plasticity optimizes sensory information processing in the olfactory bulb. *Neuron*, 91:384–396, 7 2016.

[35] Roman Shusterman, Matthew C. Smear, Alexei A. Koulakov, and Dmitry Rinberg. Precise olfactory responses tile the sniff cycle. *Nature Neuroscience*, 14:1039–1044, 8 2011.

[36] Edmund Chong, Monica Moroni, Christopher Wilson, Shy Shoham, Stefano Panzeri, and Dmitry Rinberg. Manipulating synthetic optogenetic odors reveals the coding logic of olfactory perception. *Science*, 368, 6 2020.

[37] Merav Stern, Kevin A Bolding, L F Abbott, and Kevin M Franks. A transformation from temporal to ensemble coding in a model of piriform cortex. *eLife*, 2018.

[38] Tobias Ackels, Rebecca Jordan, Andreas T. Schaefer, and Izumi Fukunaga. Respiration-locking of olfactory receptor and projection neurons in the mouse olfactory bulb and its modulation by brain state. *Frontiers in Cellular Neuroscience*, 14, 7 2020.

[39] Alex H. Williams, Ben Poole, Niru Maheswaranathan, Ashesh K. Dhawale, Tucker Fisher, Christopher D. Wilson, David H. Brann, Eric M. Trautmann, Stephen Ryu, Roman Shusterman, Dmitry Rinberg, Bence P. Olveczky, Krishna V. Shenoy, and Surya Ganguli. Discovering precise temporal patterns in large-scale neural recordings through robust and interpretable time warping. *Neuron*, 105:246–259.e8, 1 2020.

[40] Samuel Lagier, Patrizia Panzanelli, Raú E Russo, Antoine Nissant, Brice Bathellier, Marco Sassoè Pognetto, Jean-Marc Fritschy, and Pierre-Marie Lledo. Gabaergic inhibition at dendrodendritic synapses tunes oscillations in the olfactory bulb, 2007.

[41] Nathan E. Schoppa. Synchronization of olfactory bulb mitral cells by precisely timed inhibitory inputs. *Neuron*, 49:271–283, 1 2006.

[42] Kevin R. Neville and Lewis B. Haberly. Beta and gamma oscillations in the olfactory system of the urethane-anesthetized rat. *Journal of Neurophysiology*, 90:3921–3930, 12 2003.

[43] Brian Halabisky and Ben W. Strowbridge. $\gamma$ - frequency excitatory input to granule cells facilitates dendrodendritic inhibition in the rat olfactory bulb. *Journal of Neurophysiology*, 90:644–654, 8 2003.

[44] Keiji Miura, Zachary F. Mainen, and Naoshige Uchida. Odor representations in olfactory cortex: Distributed rate coding and decorrelated population activity. *Neuron*, 74:1087–1098, 6 2012.

[45] Kevin A Bolding and Kevin M Franks. Complementary codes for odor identity and intensity in olfactory cortex.

[46] Victor M. Luna and Nathan E. Schoppa. Gabaergic circuits control input-spike coupling in the piriform cortex. *Journal of Neuroscience*, 28:8851–8859, 8 2008.

[47] Agnieszk Grabska-Barwinska, Simon Barthelme, Jeff Beck, Zachary F. Mainen, Alexandre Pouget, and Peter E. Latham. A probabilistic approach to demixing odors. *Nature Neuroscience*, 20:98–106, 1 2017.

[48] Kai Shen, Sina Tootoonian, and Gilles Laurent. Encoding of mixtures in a simple olfactory system. *Neuron*, 80:1246–1262, 12 2013.

[49] Roger Ratcliff and Gail Mckoon. The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20:873–922, 2008.

[50] Zhitao Ying, Ting He, Xiaopei Wang, Wen Zheng, Ningjing Lin, Meifeng Tu, Yan Xie, Lingyan Ping, Chen Zhang, Weiping Liu, Lijuan Deng, Meng Wu, Feier Feng, Xin Leng, Tingting Du, Feifei Qi, Xuelian Hu, Yanping Ding, Xin an Lu, Yuqin Song, and Jun Zhu. Distribution of chimeric antigen receptor-modified T cells against CD19 in B-cell malignancies. *BMC Cancer*, 21, 12 2021.

[51] Kevin M. Friedman, Tracy E. Garrett, John W. Evans, Holly M. Horton, Howard J. Latimer, Stacie L. Seidel, Christopher J. Horvath, and Richard A. Morgan. Effective targeting of multiple B-cell maturation antigen-expressing hematological malignances by anti-B-cell maturation antigen chimeric antigen receptor T cells. *Human Gene Therapy*, 29:585–601, 5 2018.

[52] Sara Ghorashian, Anne Marijn Kramer, Shimobi Onuoha, Gary Wright, Jack Bartram, Rachel Richardson, Sarah J. Albon, Joan Casanovas-Company, Fernanda Castro, Bilyana Popova, Krystle Villanueva, Jenny Yeung, Winston Vetharoy, Aleks Guvenel, Patrycja A. Wawrzyniecka, Leila Mekkaoui, Gordon Weng Kit Cheung, Danielle Pinner, Jan Chu, Giovanna Lucchini, Juliana Silva, Oana Ciocarlie, Arina Lazareva, Sarah Inglott, Kimberly C. Gilmour, Gulrukh Ahsan, Mathieu Ferrari, Somayya Manzoor, Kim Champion, Tony Brooks, Andre Lopes, Allan Hackshaw, Farzin Farzaneh, Robert Chiesa, Kanchan Rao, Denise Bonney, Sujith Samarasinghe, Nicholas Goulden, Ajay Vora, Paul Veys, Rachael Hough, Robert Wynn, Martin A. Pule, and Persis J. Amrolia. Enhanced CAR T cell expansion and prolonged persistence in pediatric patients with ALL treated with a low-affinity CD19 CAR. *Nature Medicine*, 25:1408–1414, 9 2019.

[53] Michael Kalos, Bruce L. Levine, David L. Porter, Sharyn Katz, Stephan A. Grupp, Adam Bagg, and Carl H. June. T cells with chimeric antigen receptors have potent antitumor effects and can establish memory in patients with advanced leukemia. *Science Translational Medicine*, 3, 8 2011.

[54] James N. Kochenderfer, Zhiya Yu, Dorina Frasheri, Nicholas P. Restifo, and Steven A. Rosenberg. Adoptive transfer of syngeneic T cells transduced with a chimeric antigen receptor that recognizes murine CD19 can eradicate lymphoma and normal B cells. *Blood*, 116:3875–3886, 11 2010.

[55] Karen Thudium Mueller, Edward Waldron, Stephan A. Grupp, John E. Levine, Theodore W. Laetsch, Michael A. Pulsipher, Michael W. Boyer, Keith J. August, Jason Hamilton, Rakesh Awasthi, Andrew M. Stein, Denise Sickert, Abhijit Chakraborty, Bruce L. Levine, Carl H. June, Lori Tomassian, Sweta S. Shah, Mimi Leung, Tetiana Taran, Patricia A. Wood, and Shannon L. Maude. Clinical pharmacology of tisagenlecleucel in B-cell acute lymphoblastic leukemia. *Clinical Cancer Research*, 24:6175–6184, 12 2018.

[56] David L. Porter, Wei Ting Hwang, Noelle V. Frey, Simon F. Lacey, Pamela A. Shaw, Alison W. Loren, Adam Bagg, Katherine T. Marcucci, Angela Shen, Vanessa Gonzalez, David Ambrose, Stephan A. Grupp, Anne Chew, Zhaohui Zheng, Michael C. Milone, Bruce L. Levine, Jan J. Melenhorst, and Carl H. June. Chimeric antigen receptor T cells persist and induce sustained remissions in relapsed refractory chronic lymphocytic leukemia. *Science Translational Medicine*, 7, 9 2015.

[57] EMA. Assessment article (YESCARTA). 2018.

[58] Wilson W Bryan, Megan Zimmerman, Helkha Peredo-Pinto, and Xiaofei Wang. Summary basis for regulatory action - pharmacology review (Tecartus). 2020.

[59] US Food and Drug Administration. Pharmacology-toxicology review - ABECMA.

[60] US Food and Drug Administration. BLA clinical review memorandum - Kymriah.

[61] Leigh Ann Anderson. What is the cost of kymriah?, 6 2022.

[62] Can Liu, Vivaswath S. Ayyar, Xirong Zheng, Wenbo Chen, Songmao Zheng, Hardik Mody, Weirong Wang, Donald Heald, Aman P. Singh, and Yanguang Cao. Model-based cellular kinetic analysis of chimeric antigen receptor-T cells in humans. *Clinical Pharmacology and Therapeutics*, 109:716–727, 3 2021.

[63] Ahmed Salem, Ganesh Mugundu, and Aman Singh. A QSP modelling framework integrates differential kinetics of CD4/CD8 subsets and longitudinal changes in CAR-T cell immunophenotypes to characterize cellular kinetics in cancer patients., 2022.

[64] Laurence T Baxter, Hui Zhu, Daniel G Mackensen, and Rakesh K Jain. Physiologically based pharmacokinetic model for specific and nonspecific monoclonal antibodies and fragments in normal tissues and human tumor xenografts in nude mice. *Cancer Research*, 54:1517–1528, 1994.

[65] Hui Zhu, Robert J. Melder, Laurence T. Baxter, and Rakesh K. Jain. Physiologically based kinetic model of effector cell biodistribution in mammals: Implications for adoptive immunotherapy. *Cancer Research*, 56:3771–3781, 1996.

[66] Aman P. Singh and Dhaval K. Shah. A "dual" cell-level systems PK-PD model to characterize the bystander effect of ADC. *Journal of Pharmaceutical Sciences*, 108:2465–2475, 7 2019.

[67] Aman P. Singh, Xirong Zheng, Xiefan Lin-Schmidt, Wenbo Chen, Thomas J. Carpenter, Alice Zong, Weirong Wang, and Donald L. Heald. Development of a quantitative relationship between CAR-affinity, antigen abundance, tumor cell depletion and CAR-T cell expansion using a multiscale systems PK-PD model. *mAbs*, 12, 1 2020.

[68] Aman P. Singh, Wenbo Chen, Xirong Zheng, Hardik Mody, Thomas J. Carpenter, Alice Zong, and Donald L. Heald. Bench-to-bedside translation of chimeric antigen receptor (CAR) T cells using a multiscale systems pharmacokinetic-pharmacodynamic model: A case study with anti-BCMA CAR-T. *CPT: Pharmacometrics and Systems Pharmacology*, 10:362–376, 4 2021.

[69] Jim S Qin, Timothy G Johnstone, Alex Baturevych, Ronald J Hause, Seamus P Ragan, Christopher R Clouser, Jon C Jones, Rafael Ponce, Cecile M Krejsa, Ruth A Salmon, and Michael O Ports. Antitumor potency of an anti-CD19 chimeric antigen receptor T-cell therapy, lisocabtagene maraleucel in combination with ibrutinib or acalabrutinib. *J Immunotherapy*, 43:107–120, 2020.

[70] Iordanis Kesisoglou, Erfan Maddah, Ganesh M Mugundu, and Aman P. Singh. Development and application of a QSP model for cell therapies in DLBCL disease, 2022.

[71] Iraj Hosseini, Kapil Gadkar, Eric Stefanich, Chi Chung Li, Liping L. Sun, Yu Waye Chu, and Saroja Ramanujan. Mitigating the risk of cytokine release syndrome in a Phase I trial of CD20/CD3 bispecific antibody mosunetuzumab in NHL: impact of translational system modeling. *NPJ Systems Biology and Applications*, 6, 12 2020.

[72] Charles River Laboratories. NCG mouse details.

[73] Dhaval K. Shah and Alison M. Betts. Towards a platform PBPK model to characterize the plasma and tissue disposition of monoclonal antibodies in preclinical species and human. *Journal of Pharmacokinetics and Pharmacodynamics*, 39:67–86, 2 2012.

[74] Katie M Boes and Amy C Durham. Bone marrow, blood cells, and the lymphoid/lymphatic system, 2017.

[75] Anwesha Chaudhury, Andrew Stein, Stephan Grupp, John Levine, Michael Pulsipher, Doug Myers, Edward Waldron, Xu Zhu, Fraser McBlane, Rakesh Awasthi, and Edmund Waller. Conversion of cellular kinetic data for chimeric antigen receptor T-cell therapy (CAR-T) into interpretable units. 2021.

[76] Keyur Parmar, Aman P. Singh, and Ganesh M. Mugundu. Characterization of CAR-T cellular kinetics and efficacy in patients with solid tumors in presence and absence of prior lymphodepletion chemotherapy using a PBPK-PD model, 2022.

[77] Judith Feucht, Jie Sun, Justin Eyquem, Yu Jui Ho, Zeguo Zhao, Josef Leibold, Anton Dobrin, Annalisa Cabriolu, Mohamad Hamieh, and Michel Sadelain. Calibration of CAR activation potential directs alternative T cell fates and therapeutic potency. *Nature Medicine*, 25:82–88, 1 2019.

[78] E. Liu, Y. Tong, G. Dotti, H. Shaim, B. Savoldo, M. Mukherjee, J. Orange, X. Wan, X. Lu, A. Reynolds, M. Gagea, P. Banerjee, R. Cai, M. H. Bdaiwi, R. Basar, M. Muftuoglu, L. Li, D. Marin, W. Wierda, M. Keating, R. Champlin, E. Shpall, and K. Rezvani. Cord blood NK cells engineered to express IL-15 and a CD19-targeted CAR show long-term persistence and potent antitumor activity. *Leukemia*, 32:520–531, 2 2018.

[79] M. Spiridon and W. Gerstner. Effect of lateral connections on the accuracy of the population code for a network of spiking neurons. *Network: Computation in Neural Systems*, 12:409–421, 1 2001.

APPENDIX A

# Supplementary Analyses

## A.1. Poisson spike trains' variance of rates decreases linearly with $T$

A Poisson spike train that has (on average) $\lambda$ spikes per unit time has a variance that scales linearly with time $T$. Its variance of rates scales with $1/T$.

$$Var\left[\frac{1}{T}\text{ Spikes in time } T\right] = \frac{1}{T^2}Var\left[\text{ Spikes in time } T\right] \tag{A.1}$$
$$\propto \frac{1}{T^2}T$$
$$\propto \frac{1}{T}$$

## A.2. Overlapping measurement windows do not substantially change estimates of $\mathcal{F}_{opt}$ at sufficiently large time series durations.

In steady-state analyses, we assume that mitral cell spike data are stationary processes and place measurement windows randomly throughout the time series. This allows the possibility of different samples to overlap and thus not be independent.

To address the potential issue of overlapping data samples overestimating correlations, we test several measurement durations and time series durations to assess how $\mathcal{F}_{opt}$ is impacted by overlaps. In each of these measurements, $\Delta\vec{\mu}$ is held constant, so any estimation variability solely arises from estimations of the covariance terms. Indeed, as the measurement duration becomes larger, it takes a progressively longer time-series in

order to eliminate any confounding influence of overlapping bins (Figure A.1). In our analyses, we use time series lengths of at least 10,000 ms and measurement durations of at most $T = 200$ ms. This parameter combination exists in a range that does not appear to be impacted by overestimation.

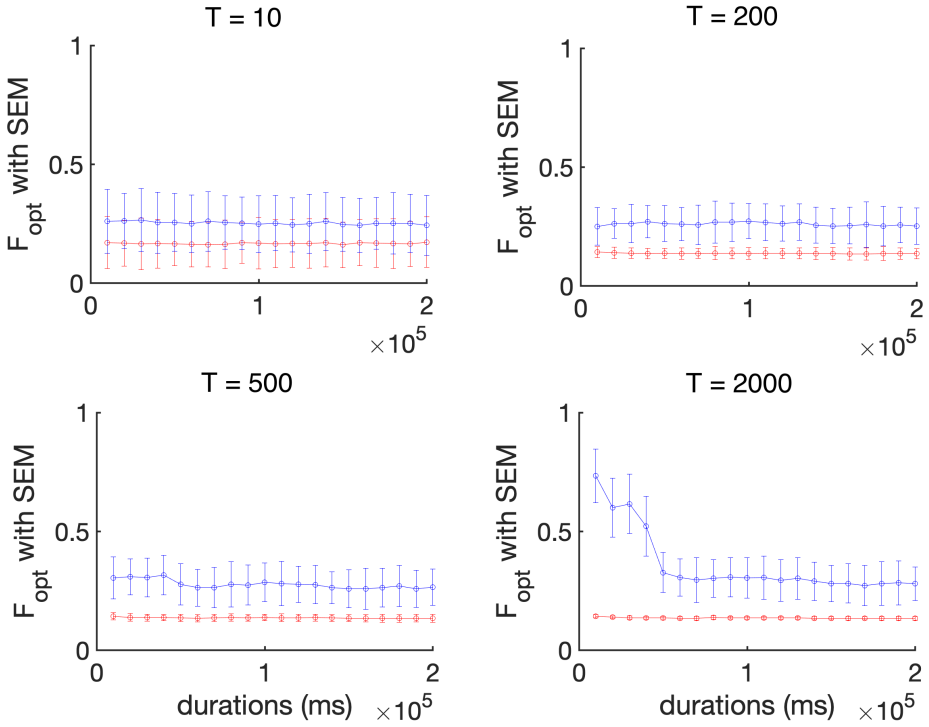


Figure A.1. Estimates $\mathcal{F}_{opt}$ (blue lines) with standard errors of the mean across samples. Red lines indicate measurements of $\mathcal{F}_{uncorr}$. These measurements were done in a globally-connected network.

## A.3. Infinitely fast granule cells strongly reduce variance single-connected networks.

In analyzing the single reciprocal network, we observed that the synaptic response time of the GC affects network oscillations. In this 1 MC-GC pair, the mitral cell spikes and increments its $\bar{x}$ variable, which increments the GC's evolutionary equation, and when

the GC spikes, it increments its own $\bar{x}$ variable, which increments the MC's evolutionary equation. During one alternative implementation, the MC spikes, increments its own $\bar{x}$ variable, and the value of that variable is fed back into the MC's inhibitory signal within the same time iteration. In other words, the MCs receives instant inhibitory feedback each time it spikes.

Firing rate decreases with strength of inhibition $w$ similarly as before (Figure A.2a), but variance decreases much more strongly with $w$ (at large window sizes), lacking the characteristic upswing that occurs at large $w$ (Figure A.2b). Additionally, plots of scaled variance with $T$ shows a stronger downward trend, both at $w = 2$ (Figure A.2c) and $w = 50$ (Figure A.2d). In summary, the initial upshoot at small $T$ (reflective of the gamma frequency) vanishes, but the noise-shaping at larger measurement windows remains.

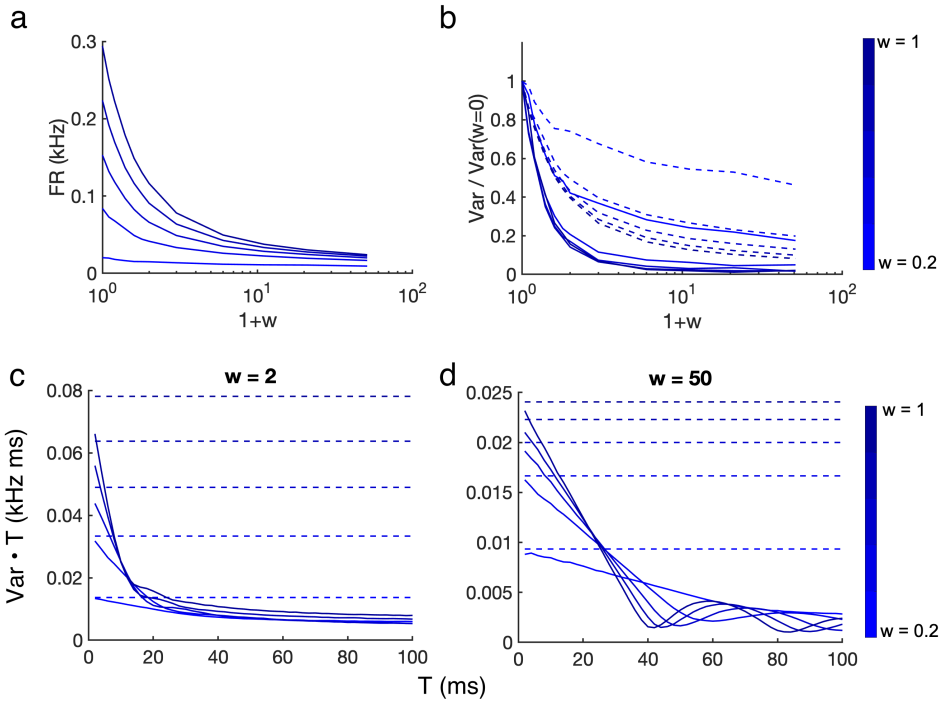


Figure A.2. Single MC-GC reciprocally connected network with infinitely fast granule cell. (a) Firing rate with $w$ (b) Scaled variance as a proportion of $w = 0$, at $T = 100$. (c) Variance scaled with $T$. Dashed lines indicate variance of the corresponding Poisson network. (d) Same as (c) but for $w = 50$. Different shades of blue indicate relative strengths of OSN input.

## A.4. Neurons with shared inhibition are less correlated to their granule cells and experience less variance reduction.

Analyses of how network size impacts noise correlations originally arose from an observation of globally-connected networks. Discriminability trends were qualitatively different for networks of 2 MC and 10 MC, dashing our initial expectations of extrapolating results of a 2-MC network to larger networks (Figure A.3). In the smaller network, $\mathcal{F}_{uncorr}$, which ignores pairwise covariances, substantially improved with small amounts of inhibition, suggesting that in those regimes, individual cell variance was reduced faster than the corresponding squares of mean differences. This finding however did not extrapolate to a network of 10 MC, in which both $\mathcal{F}_{opt}$ and $\mathcal{F}_{uncorr}$ barely improved at small inhibition and mostly decreased with inhibition.

$\mathcal{F}_{uncorr}$ assumes each MC is independent and is linear in the number of MCs. For the same difference in mean rates, if the cells in 2-MC and 10-MC networks had the same variances, we'd expect exactly a 5-fold increase in $\mathcal{F}_{uncorr}$ (in general, if the system size increases by a factor of $K$, uncorrelated discriminability increases by the same factor of $K$). Instead, we find that for any amount of nonzero inhibition, the variance per cell in a 10-MC network was larger than that of a 2-MC network, and that the scaling factor was not the same for all $w$.

## A.5. Contributions of $\Sigma$ terms in $\mathcal{F}_{opt}$ and $\mathcal{F}_{uncorr}$.

This section shows how the $\Sigma$-terms $a$, $b$, and $c$ shape $\mathcal{F}_{opt}$ and $\mathcal{F}_{uncorr}$ in a globally connected network exposed to symmetric block stimuli.

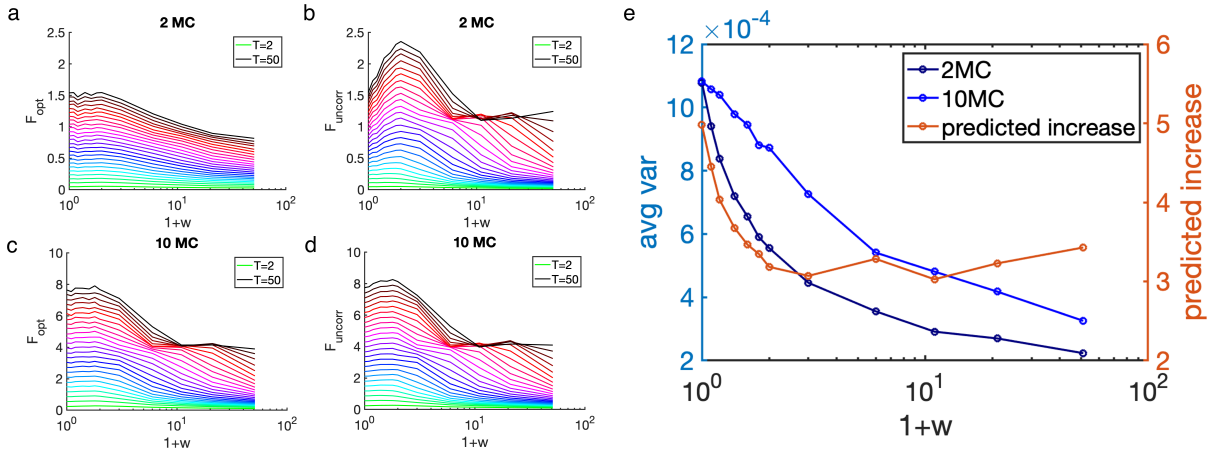


Figure A.3. Discriminability in 2-MC and 10-MC networks, **not** scaled with $T$. (a) $\mathcal{F}_{opt}$ with $1+w$, in a 2-MC network at different measurement durations from $T$ = 2 to 50. (b) Corresponding $\mathcal{F}_{uncorr}$ in a 2-MC network. (c) and (d) Same as (a) and (b) but for a 10-MC network. (e) Left $y-$axis: Average variance for a 2-MC (dark blue) and 10-MC (royal blue) globally-connected network at inhibitory coupling y the number of cells, resulting in networks whose cells fired at the same average rates. Qualitative differences in $\mathcal{F}_{uncorr}$ suggests differences in their variances directly resulting from changing the network size.

Since $b$ and $c$ each have a multiplicative factor that scales with the system size, their contributions are usually larger than that of $a$ (Figure A.4, top row). However, since $b$ and $c$ show similar qualitative trends but with opposite signs, their contributions mostly cancel out, and the cumulative contribution from the covariance is not substantially different from the contribution of just the variance term.

Taking the inverses of variance terms highlights small differences in qualitative behavior (Figure A.4 middle row). We find that the contribution of variance and covariance to discriminability trends upwards as inhibition increases, suggesting that recurrent inhibition helps modulate the noise in MC activities. Despite the general trend upwards, these covariance contributions are not monotonic with window size, again reflecting the

rhythmicity of the system and the importance of measurement size relative to the period of network oscillations.

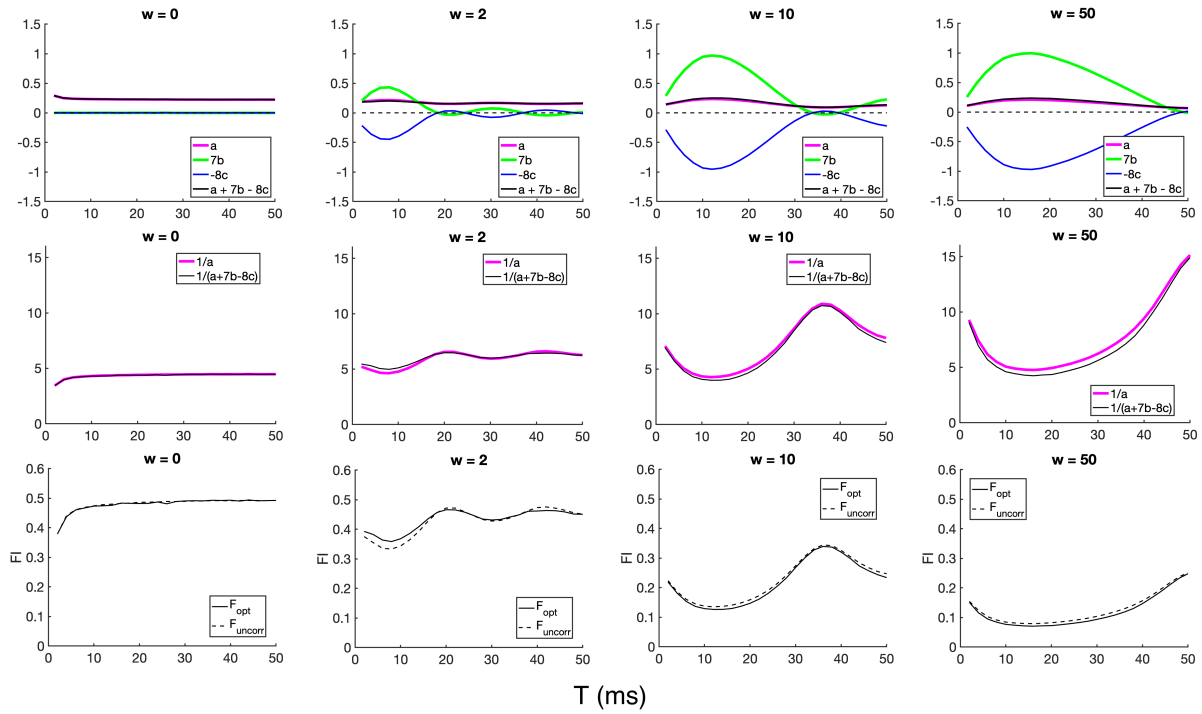


Figure A.4. Effect of covariance terms in computing linear discrimination. (a) Terms of the covariance matrix with window size. Solid black line is the sum. Dashed black line indicates zero. (b) Variance contribution (pink line) and variance contribution including off-diagonal terms (black line). (c) Optimal Fisher information (solid black line) and Fisher Information without correlations (dashed black line), scaled with $T$.

## A.6. Pairwise covariances of MCs exposed to high-contrast blocks

Many of our networks are exposed to block stimuli defined by two levels of input. There are correspondingly three possible combinations of pairwise covariances: between mitral cells both receiving the higher input, between mitral cells both receiving the lower

input, and between mitral cells receiving one of each input. For notational convenience, denote these quantities $Cov_{high}$, $Cov_{low}$, and $Cov_{high,low}$.

For all pairs of block stimuli, $Cov_{high,low}$ is closer to $Cov_{low}$ than $Cov_{high}$, and it is closer yet in high-contrast blocks. To see why, we more closely examine the network at $w$ = 50, an inhibitory strength that is strong enough to produce system rhythms, and measurement durations $T$ = 16 ms and 50 ms. The smaller bin $T$ = 16 corresponds to a window size near the maximum of the covariance curves, whereas $T$ = 50 corresponds to a window size near the system period, at which all pairwise covariances are near zero.

In high-contrast blocks, cells receiving stronger input spike more frequently in between inhibitory cycles than cells receiving weaker input (Figure A.5abc). In their power spectral densities, MCs receiving weaker stimulation also have lower amplitude peaks (Figure A.5bc), even though their peaks align to the same frequencies as their higher-stimulated counterparts. Lower-amplitude-peaks in the PSD indicate that these cells have less variance content (Appendix B.5.2).

At $T$ = 16 ms, the majority of randomly placed measurements pick up periods of silence. Relative frequencies of spike counts have more distributional weight near zero, with some weight at higher numbers. The stronger the stimulus intensity in those cells, the larger weight is placed at higher spike counts (Figure A.5d), which also results in marginally higher mean spike count, as well as higher signed deviations from the mean in both directions (Figure A.5e). These slightly-higher deviations correspond to a higher covariance; the numerical covariance is equal to the sample mean of products of deviations. Pairs of highly-stimulated cells see occasional large deviations from their own mean, which increases their pairwise covariance.

These discrepancies vanish at $T = 50$ ms. At this larger bin, windows capture a full period of activity, and even though spike count histograms vary between high- and low-stimulated cells (Figure A.5g), measurements each contain similar deviations from their own means (Figure A.5h), and have similar, near-zero covariances (Figure A.5i).

In contrast, activity from a network receiving low-contrast blocks input show much smaller differences between differently-stimulated cells. Their raster plots are more visually similar (Figure A.6a), their corresponding power spectral densities are similar, and their spike-counts histograms are shifted similarly to that of the high-contrast case, but to a lesser degree (Figure A.6bc). Whereas different-stimulated cells in high-contrast blocks had a much lower pairwise covariance than high-stimulated cells at small window sizes, we see no such discrepancy here. Even at small window sizes, spike counts across samples follow similar distributions, with neither category having a much stronger right tail than the other. Their deviations from their respective means also follow similar distributional shapes (Figure A.6de). As a result, pairwise products of deviations are all similar to each other, independent of the level of stimulation each cell in the pair receives (Figure A.6f). As in high-contrast blocks, we find that at a sufficiently large window size that captures a full period of activity, measurement deviations from their means are clustered near zero, and their products are all close to each other (Figure A.6ghi).

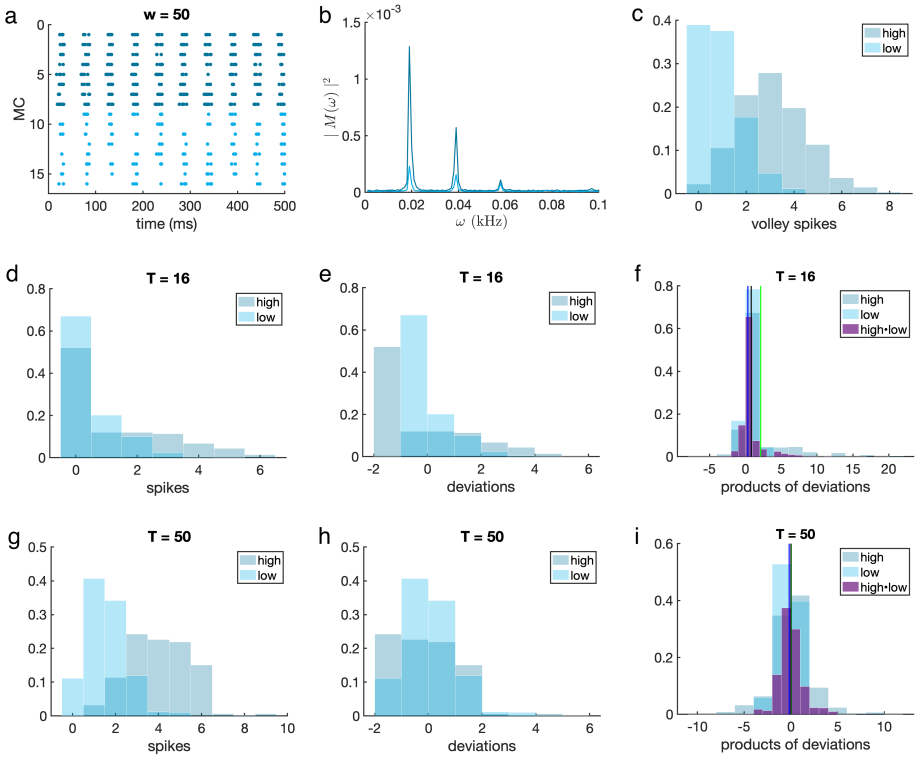


Figure A.5. Spikes, power spectral densities, and spike statistics for a network exposed to high-contrast blocks. (a) Raster plots of 16 MCs in a 500 ms interval. Darker teal corresponds to cells receiving higher inputs; lighter teal corresponds to cells receiving lower inputs. (b) Power spectral density for higher-stimulated and lower-stimulated cells. PSD is averaged over all cells in the high-stimulated and low-stimulated groups, respectively. (c) Relative frequency of spike counts in each periodic volley. (d) Relative frequencies of spike counts captured in randomly placed windows of measurement duration $T = 16$. (e) Deviations of (d) from their means. (f) relative frequencies of products of deviations for high-stimulated cells (green), low-stimulated cells (blue), and differently-stimulated cells (purple). Vertical black line indicates sample mean of different-stimulated products of deviations, vertical blue line corresponds to low-stimulated pairs, and vertical green line corresponds to high-stimulated pairs. (g-i) Same as (d–f) but for $T = 50$ ms.

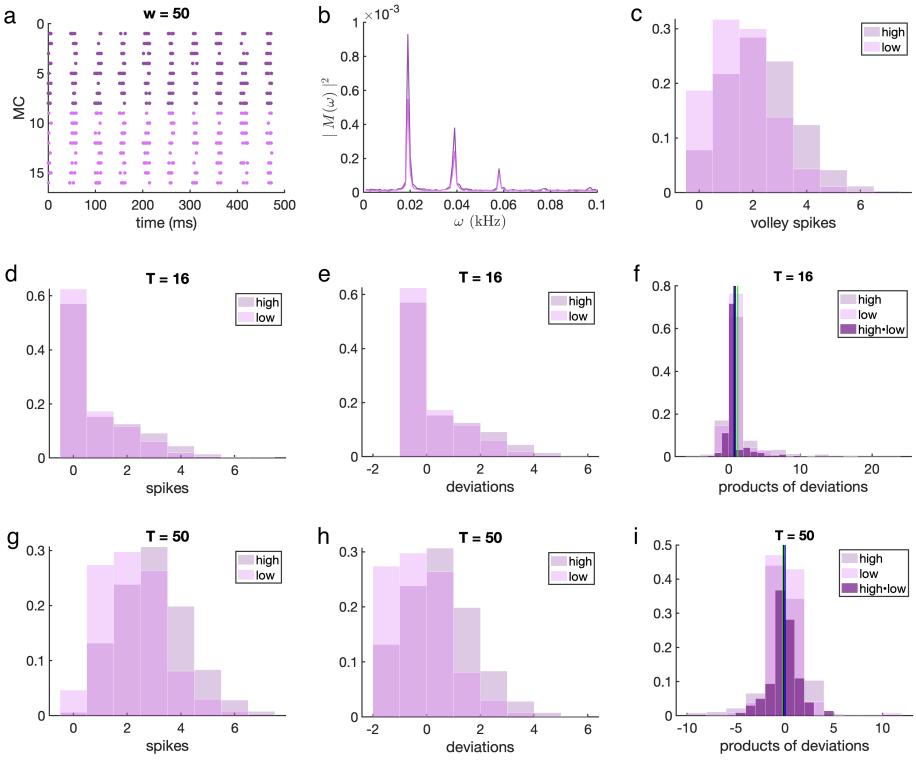


Figure A.6. Same as for Figure A.5 but for low-contrast blocks.

## A.7. Stimuli with overlapping set of MCs.

These analyses were developed as a precursor to skewed Gaussians, in order to better examine how selectively suppressing distracting MCs impact discriminability. For each of high-contrast blocks and low-contrast blocks, we add a set of eight cells that fire at the same intensity across both stimuli in a pair. We term this middle set the "distractor set" and cells outside this group as MCs in the "wings." (Figure A.7).

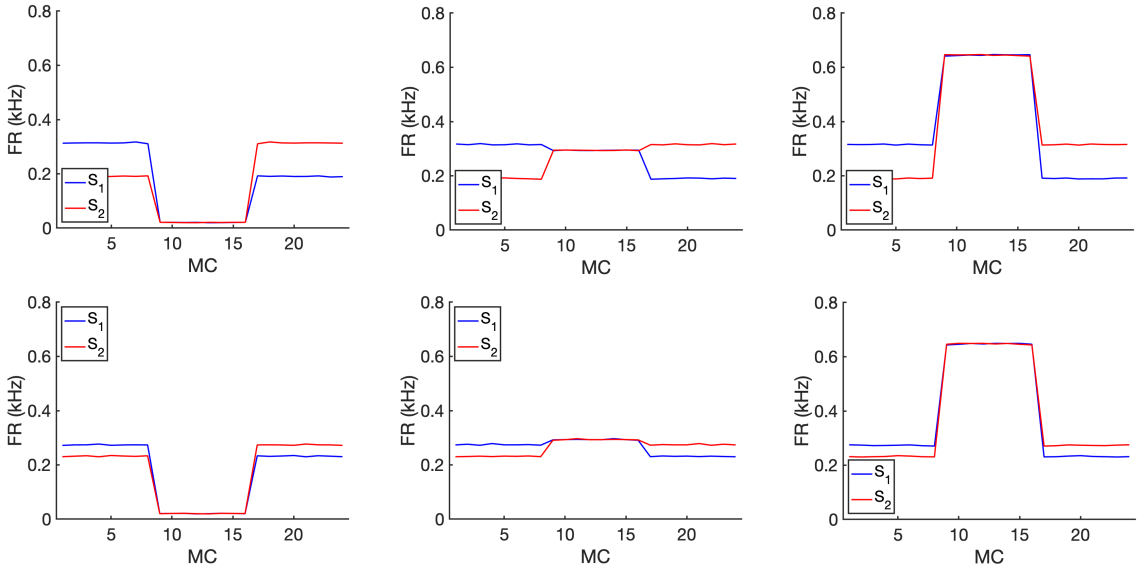


Figure A.7. High-contrast blocks (top row) and low-contrast blocks (bottom row) stimuli with a third set of firing cells driven by (left to right) 0.2, 1, and 2 Hz inputs.

Note that the distractor set does not impact discriminability in a single-connected network, since $\Delta\vec{\mu}$ is zero in entries corresponding to distractor MCs, and noise correlations between cells in the wings are not impacted by the activity of the distractor.

### A.7.1. Distractor set activity dominates the system rhythm.

In a globally-connected network, distractor activity impacts noise correlations within cells through shared inhibition, which may impact discriminability. As the activity of the distractor increases, the overall frequency of the system increases (Figure A.8ae). When there is low distractor activity, granule cells are driven primarily by cells in the wings, producing a rhythm (Figure A.8bf). As amount of distractor activity increases, these middle blocks dominate the overall MC response, driving the GCs more strongly and and increasing the frequency of the system rhythm (Figure A.8ae).

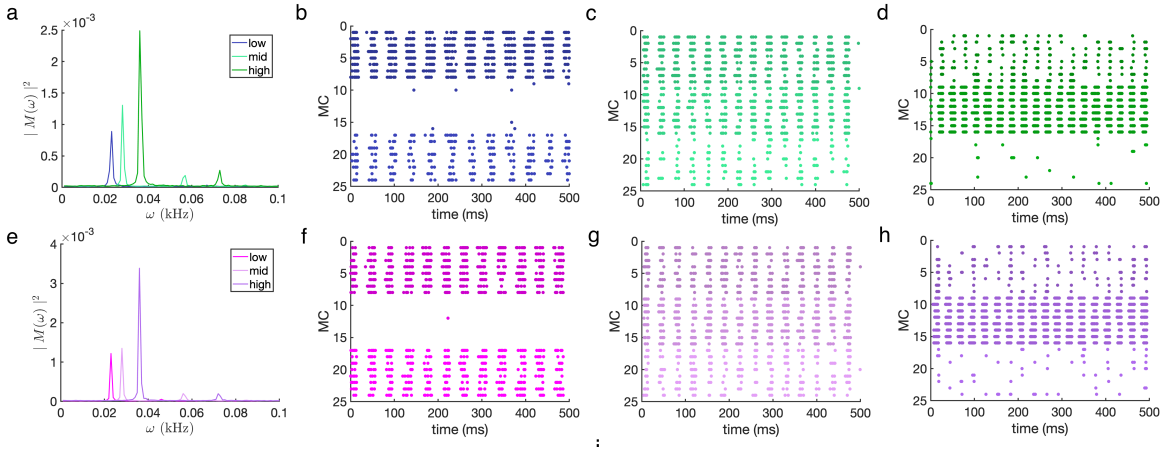


Figure A.8. Globally-connected network exposed to stimuli with distractor set. (a) Power spectral density for high-contrast blocks with distractor set. Increasingly darker shades of green correspond to increasing levels of activity from this middle set. PSD is averaged over all 24 cells in the network. (b) Raster plot for low levels of distractor activity. (c) Raster plot for middle levels of distractor activity (d) Raster plot for high levels of distractor activity. (e-h) Same as (a-d) for low-contrast blocks.

### A.7.2. Noise correlations are reduced by strongly stimulated distractor set.

Scaled variance qualitatively behaves similarly to that of stimuli without the distractor (Figure A.9, A.10). As the cells in the middle block are driven more strongly (light green to dark green lines in the below figure), not only does the overall variance decrease, but the first local minimum shifts to shorter measurement durations. Higher activity from the distractor block drives the granule cells more strongly, creating a feedback rhythm with shorter volleys and faster recoveries. In other words, the distractor entrains the network at high levels of activity.

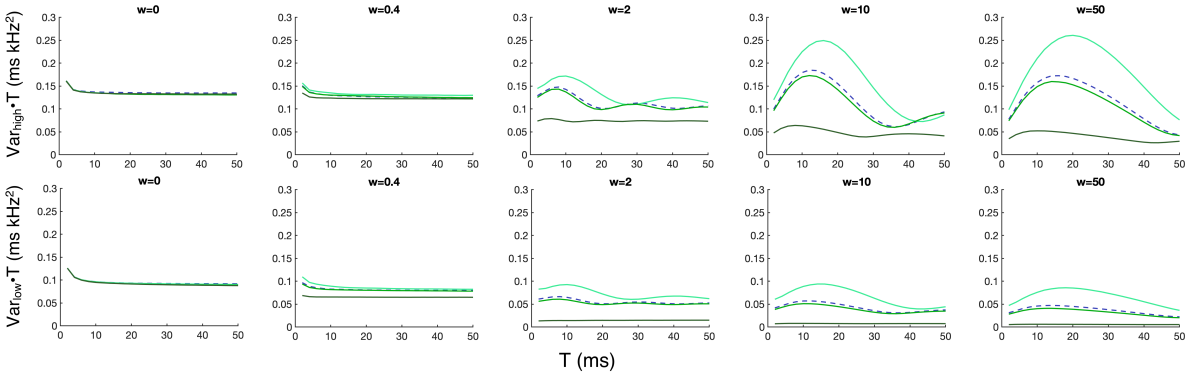


Figure A.9. Variances in a globally-connected network exposed to high-contrast blocks with distractor set: scaled variance of cells receiving the higher stimulus (top row) and smaller stimulus (bottom row). Increasingly darker shades of green correspond to increasing activity of the distractor cells. Dashed line is the corresponding variance measurement in the absence of distractor blocks.

Pairwise covariances follow similar qualitative trends as variance terms. As the activity of the distractor set increases, covariance decreases, and the measurement duration of the first local minimum decreases. When the activity of the middle block is very large, all variance and covariance terms become very small (Figures A.11, A.12).

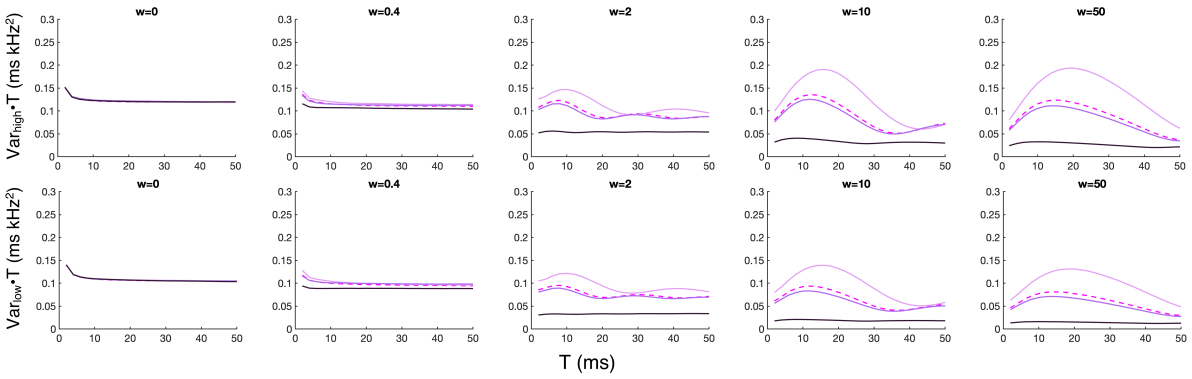


Figure A.10. Same as for Figure A.9 but for low-contrast blocks.

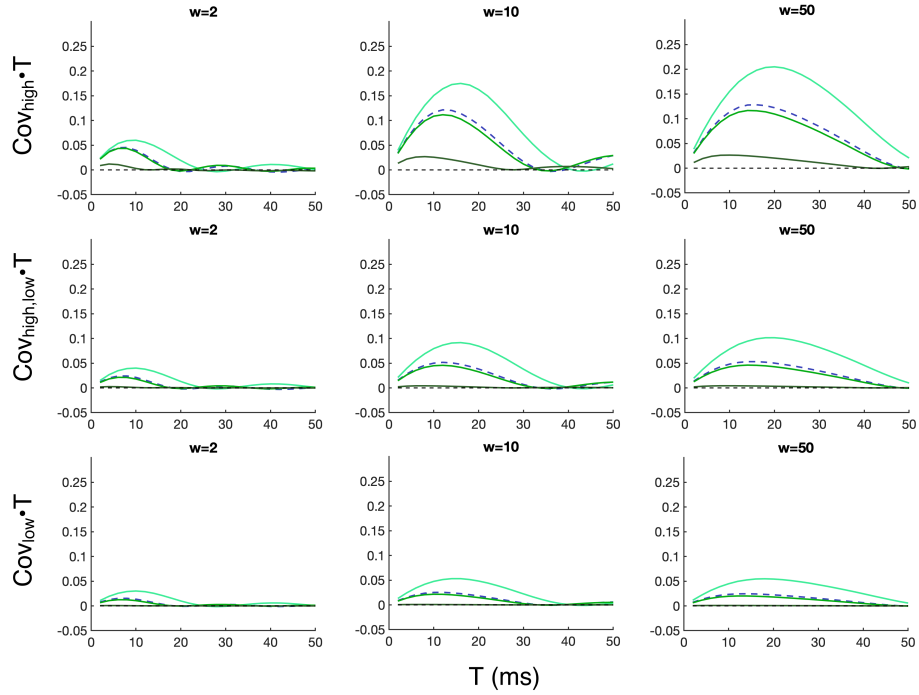


Figure A.11. High-contrast blocks with distractor set: scaled pairwise covariance of (top row) cells both receiving the high stimulus, (middle row) cells receiving different inputs, and (bottom row) cells receiving the lower stimulus. Increasing darker shades of green correspond to increasing activity of the distractor set. Dashed line is the corresponding variance measurement in the absence of distractor blocks.

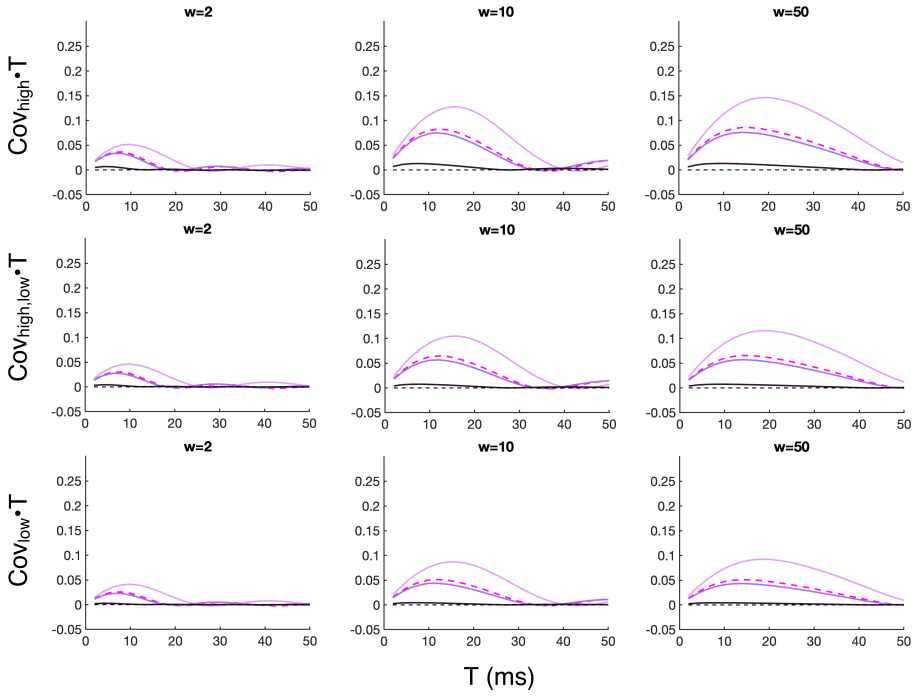


Figure A.12. Same as for A.11 but for low-contrast blocks. Increasing darker shades of purple correspond to low, medium, and high activity of the distractor set.

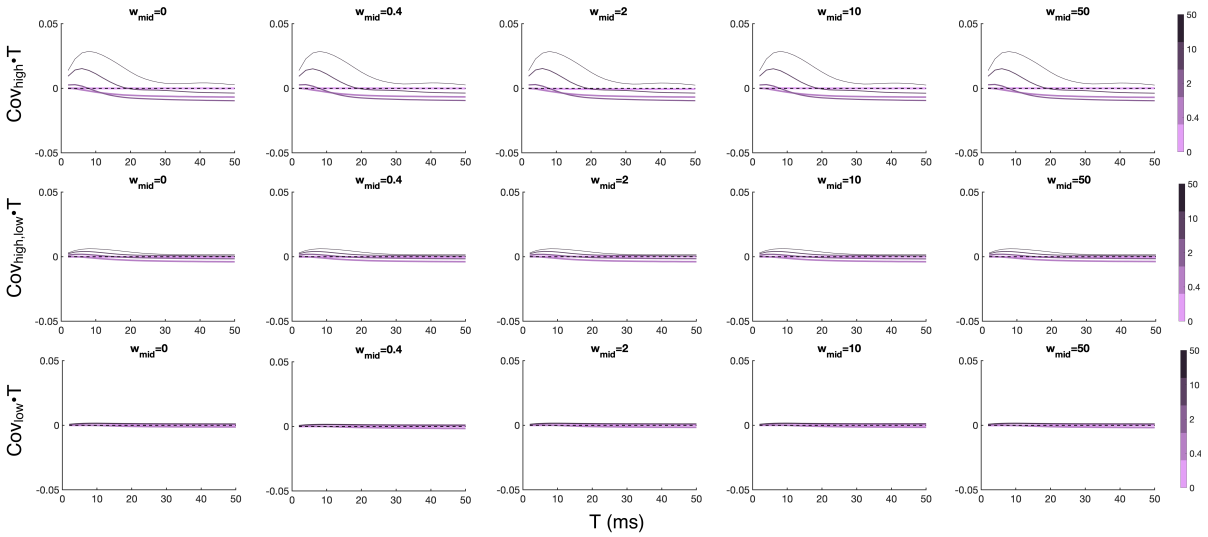


Figure A.13. Pairwise covariance of cells scaled by $T$ between MCs both receiving the higher stimulation (top row), MCs receiving different stimulation (middle row), and both cells receive the lower stimulation (bottom row). Left to right indicate different levels of preferential middle stimulation. Increasing darker shades of purple indicate different levels of within-wing stimulation. Dashed black line is zero.

### A.8. Distractor set activity impacts whether covariances are beneficial

This section shows that the activity of the distractor set can change whether pairwise covariances benefit discriminability in globally-connected networks.

In the presence of a distractor set of MCs whose firing rates are equal across both stimuli, $\Delta\vec{\mu}$ is zero in those entries and only nonzero in the wings. As a result, the quantity $\mathcal{F}_{opt} = \Delta\vec{\mu}^T\Sigma\Delta\vec{\mu}$ is governed only by the contributions from MCs in the wings. While in the previous globally-connected regime, the pairwise covariances of cells in the wings depend on activity contrasts, that is not the case here. We show a few examples below.

Our previous analyses of low-contrast blocks showed that pairwise covariance is similar regardless of whether two cells are high-, low-, or differently-stimulated. This results in values of $b$ that are similar to values of $c$, and correspondingly, $\left(\frac{Nm}{2} - 1\right) b < \frac{Nm}{2} c$. Linear discriminability is improved when these offdiagonal terms are taken into account. The ratio $\mathcal{F}_{opt}$ to $\mathcal{F}_{uncorr}$ is greater than one, suggesting that it is beneficial to include pairwise covariances. However, if the activity of the middle set is sufficiently high, then pairwise covariance of cells in the wings is pushed to near-zero levels. At these activity levels, pairwise covariance does not substantially affect linear discriminability, and $\mathcal{F}_{opt}/\mathcal{F}_{uncorr}$ is nearly one (Figure A.14c).

In networks exposed to high-contrast blocks, the covariances of similarly-stimuated cells was substantially higher than that of differently-stimulated cells. In this scenario, $\left(\frac{Nm}{2} - 1\right) b > \frac{Nm}{2} c$ and these off-diagonal covariance terms are detrimental to linear discriminability, particularly at smaller measurement bins. As the activity of the distractor set increases, the relative detriment of the covariance term also increases, resulting in

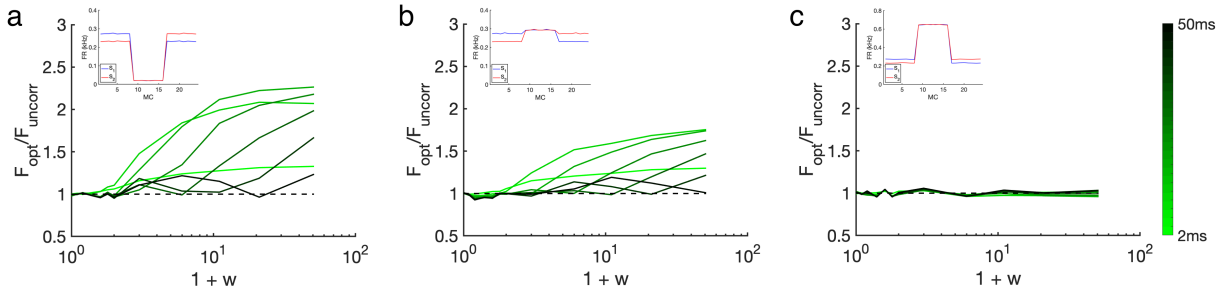


Figure A.14. $\mathcal{F}_{opt}/\mathcal{F}_{uncorr}$ for low-contrast blocks with (a) low (b) middle and (c) high amounts of distractor activity. Insets show average MC activity taken over the entire time series. Vertical axis limits are different in (c) to show height of distractor block activity. Shades of increasingly dark green indicate increasing measurement bin. Discriminability is scaled with $T$.

an $\mathcal{F}_{opt}/\mathcal{F}_{uncorr}$ ratio that is below one and decreases further with the strength of the distractor set (Figure A.15).

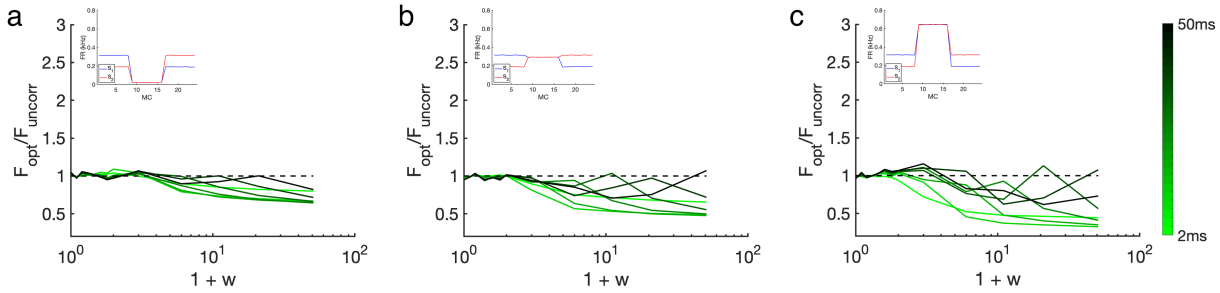


Figure A.15. Same as for Figure A.14 but for High-contrast blocks.

### A.8.1. Inhibition impacts noise correlations only in the cells it suppresses.

The following results are not surprising, but we examine this case as a building block towards understanding how stimulus-based inhibition impacts network performance in response to skewed Gaussians. Specifically, we consider a network exposed to stimuli containing a distractor set, but there is no shared connectivity between the distractor set

and cells in the wings. This is functionally equivalent to a system with two independent subnetworks, each with their own global inhibition.

Inhibition within each subnetwork selectively decreases firing rates of cells within that region (Figure A.16). Similarly, inhibition only reduces noise correlations between the cells it suppresses (see Appendix A.7.2 for all figures of noise correlations).

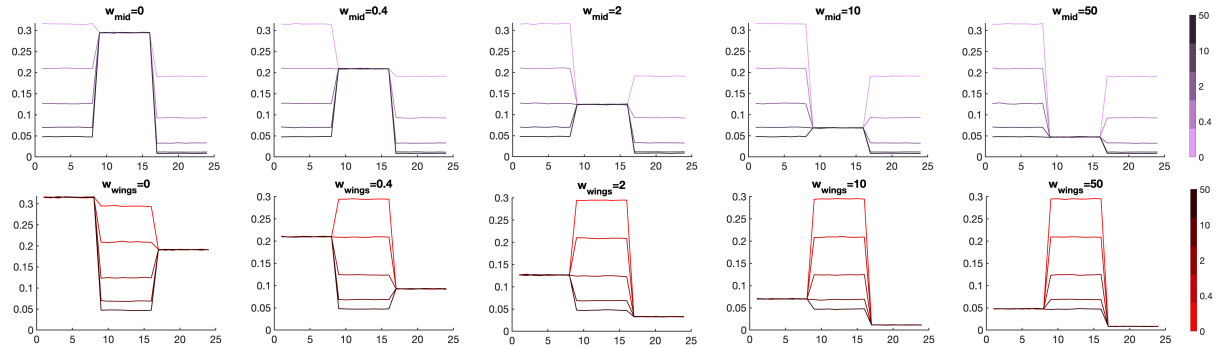


Figure A.16. (Top row) Average firing rates of 24 MCs. Left to right: increasing middle-block inhibition. Increasing darker shades of purple indicate increasing within-wing stimulation. (Bottom row) Left to right: increasing within-wing inhibition. Increasingly darker shades of red correspond to increasing middle block inhibitions.

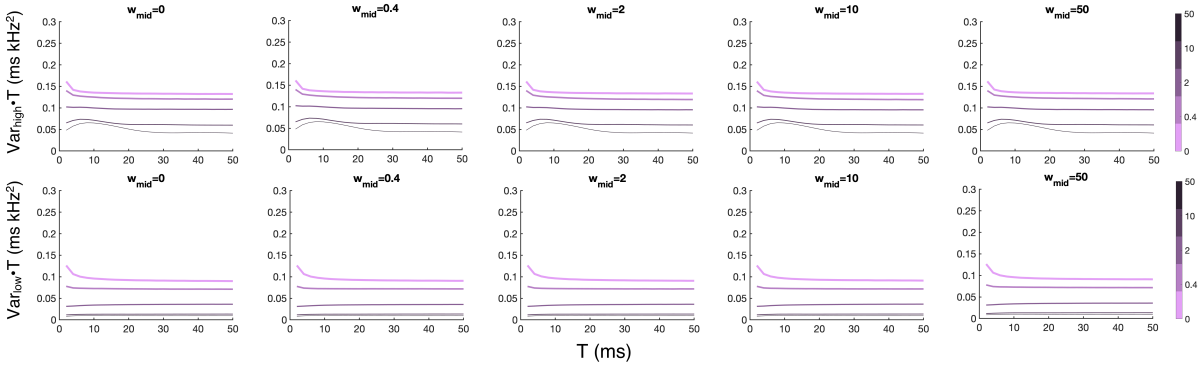


Figure A.17. Scaled variance of cells receiving the higher stimulation (top row) and the lower stimulation (bottom row). Left to right indicate increasing levels of preferential middle stimulation. Increasing darkershades of purple indicate different levels of within-wing stimulation.

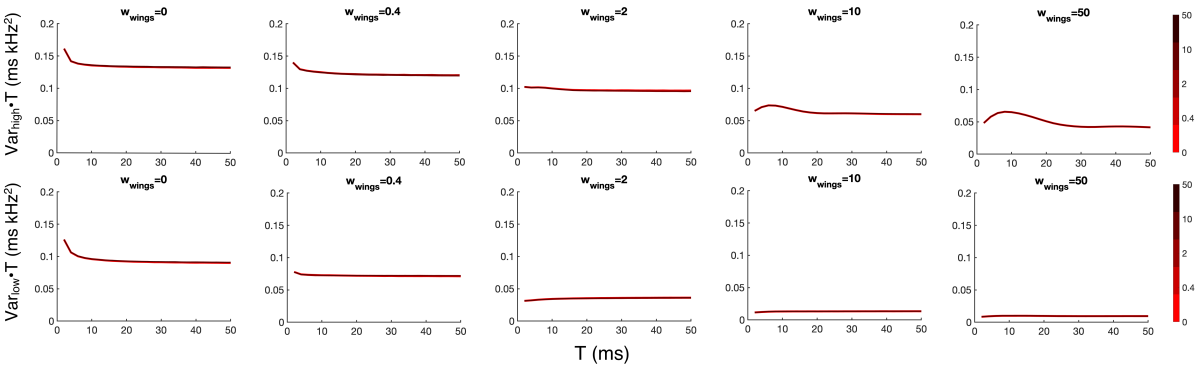


Figure A.18. Scaled variance of cells receiving the higher stimulation (top row) and the lower stimulation (bottom row). Left to right indicate increasing levels of preferential within-wing stimulation. Increasingly darker shades of red indicate different levels of middle stimulation. Lighter shades of red are plotted at larger linewidths for visibility.

### A.9. Random linear discriminability accounts for noise in all directions.

This section serves to visualize how $\mathcal{F}_{random}$ captures noise correlations, both in the optimal and nonoptimal directions.

We consider a toy example with two mitral cells and observe how $\mathcal{F}_{opt}$ is affected by noise-shaping. As a baseline, we simulate noisy spiking data of two mitral cells that are uncorrelated with each other but have equal variance distributions (Figure A.19a, top row). We also consider a case where the signs of the correlations are the same but larger in each direction (Figure A.19a, bottom row).

The distribution of $\mathcal{F}_{random}$ has a support whose lower limit is zero and upper limit is $\mathcal{F}_{opt}$. When the cells are negatively correlated, $\mathcal{F}_{random}$ has more distributional weight towards higher values. Conversely, when cells are pairwise positively correlated, $\mathcal{F}_{random}$ becomes skewed right, indicating a higher prevalence of worse discriminability. In general, the size of the odd eigenvalue controls $\mathcal{F}_{opt}$ and thus the upper limit of $\mathcal{F}_{random}$, while the relative size of the odd and even eigenvalues determines the distribution skew.

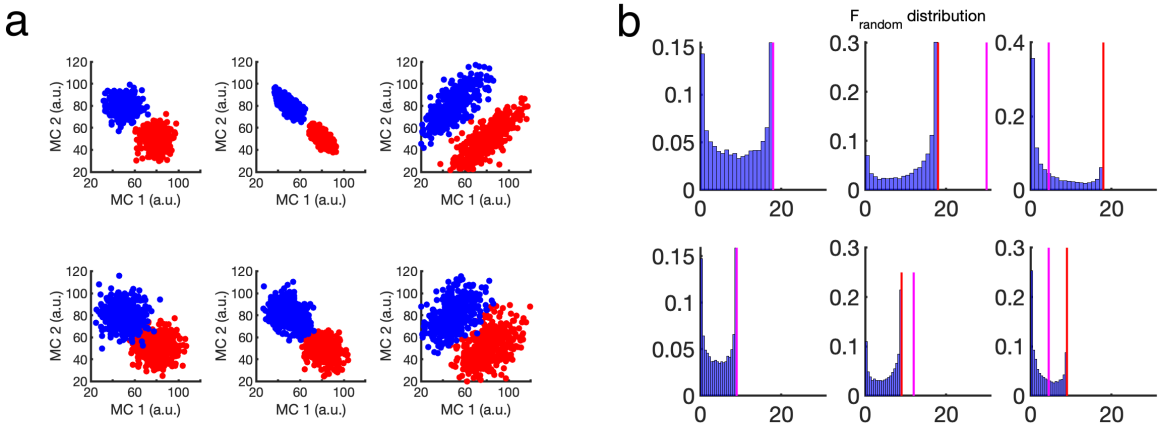


Figure A.19. (a) Firing rates (arbitrary units) of two MCs in response to symmetric stimuli (red and blue). In each of the top and bottom rows, the odd eigenvalue of the covariance sum is equal across all three examples. In the bottom middle, the even eigenvalue is equal in magnitude to the odd eigenvalue of each example in the top row. (b) Corresponding probability distributions of $\mathcal{F}_{random}$ for the six cases above. Red vertical line indicate $\mathcal{F}_{opt}$ for each dataset. Magenta vertical line indicates $\mathcal{F}_{uncorr}$ for each dataset.

## A.10. On/off sniff: average firing rates and discriminability plots

This section contains heatmaps of average MC firing rate and discriminability plots for all networks simulated with a Step function inhalation.

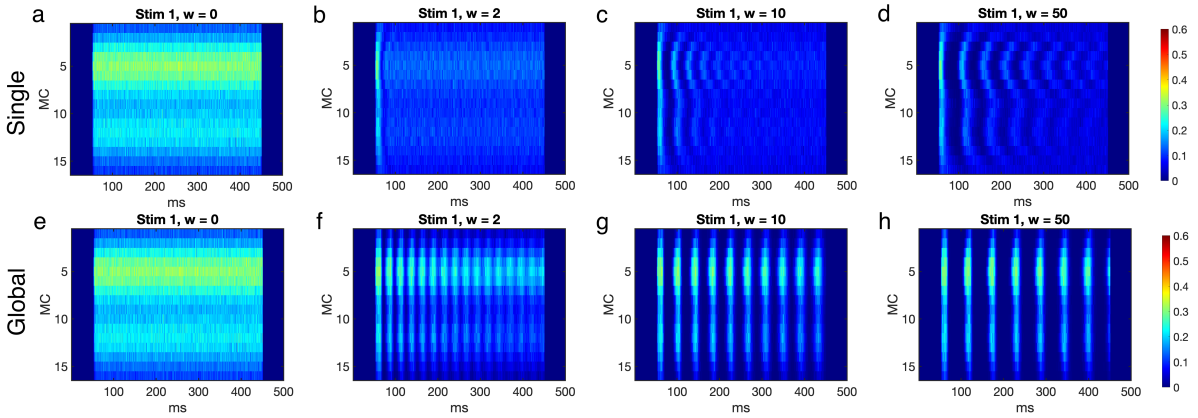


Figure A.20. Firing rates of a network exposed to Gaussian mixtures averaged over 400 sniff cycles for (top row) Single-connected and (bottom row) Globally-connected networks, for (ac) $w = 0$, (bd) $w = 2$, (ce) $w = 10$ and (dh) $w = 50$.

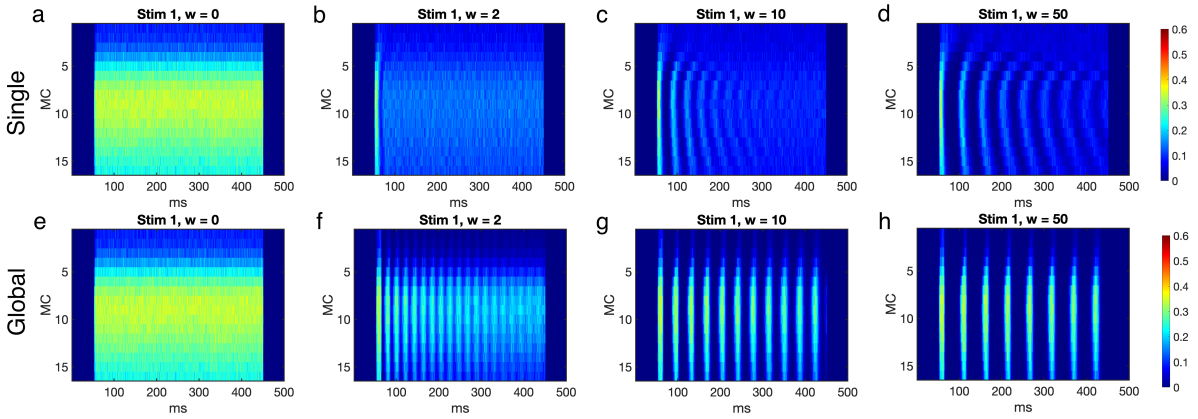


Figure A.21. Firing rates of a network exposed to skewed Gaussians averaged over 400 sniff cycles for (top row) Single-connected and (bottom row) Globally-connected networks, for (ac) $w = 0$, (bd) $w = 2$, (ce) $w = 10$ and (dh) $w = 50$.

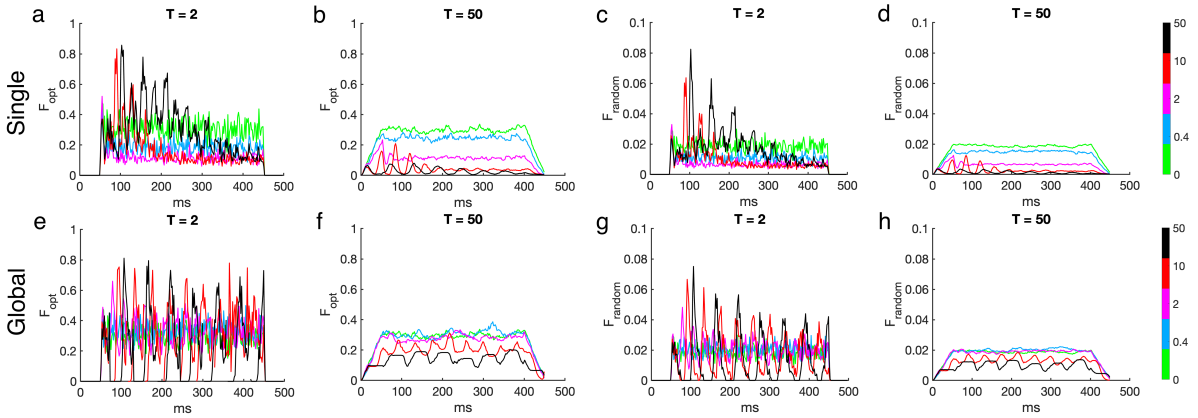


Figure A.22. $\mathcal{F}_{opt}$ and $\mathcal{F}_{random}$ for single-connected (top row) and globally connected (bottom row) networks exposed to Gaussian mixtures. Data is scaled with $T$.

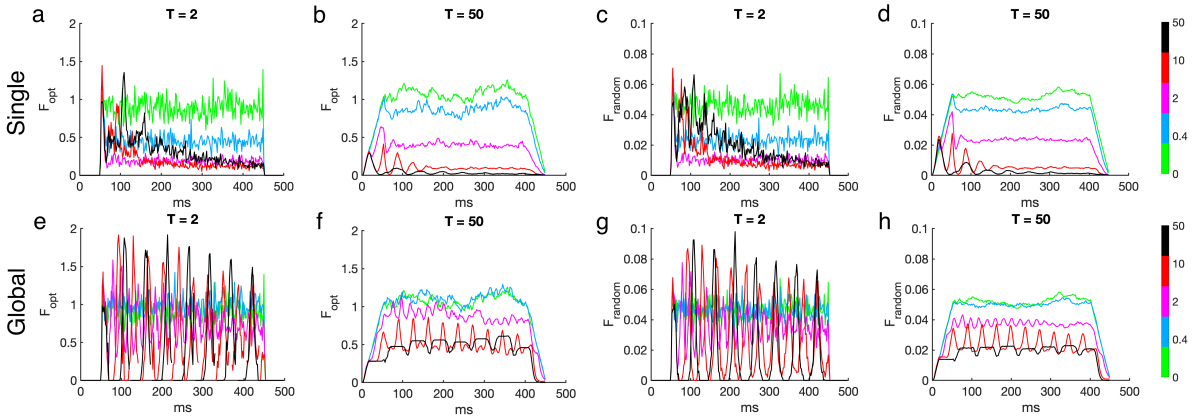


Figure A.23. $\mathcal{F}_{opt}$ and $\mathcal{F}_{random}$ for single-connected (top row) and globally connected (bottom row) networks exposed to skewed Gaussians. Data is scaled with $T$.

Figures A.20 and A.21 show the average sniff activity for Gaussian mixtures and skewed Gaussians, respectively. In a single-connected network, sufficiently large inhibition induces separate rhythms for each MC, with the most strongly stimulated cells spiking earlier in each wave. Bands of activity lose definition over time because MCs excite their GCs at variable times, with most highly excited MCs incurring inhibition earlier. Global connected networks demonstrate much more regular bands of spiking that persist

throughout the sniff; in these networks, GC inhibition is delivered synchronously for all MCs.

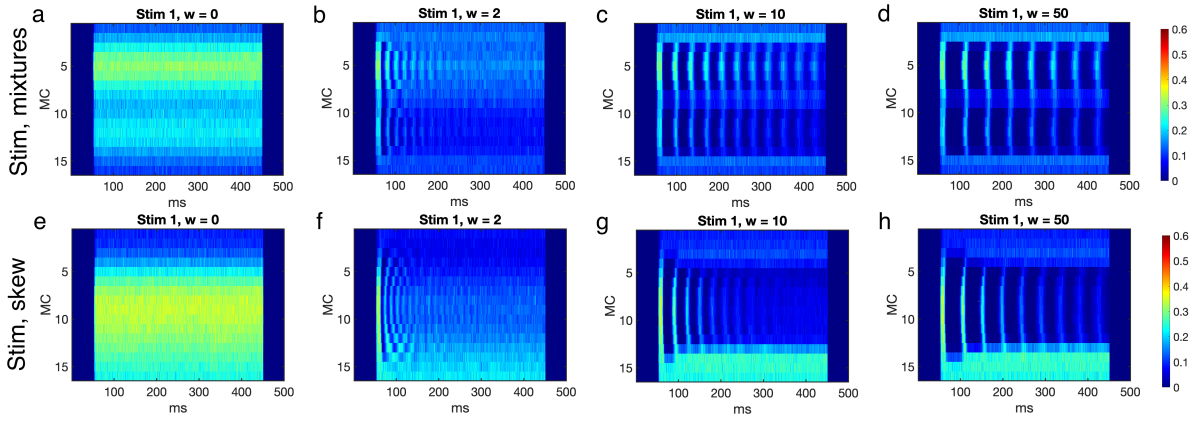


Figure A.24. Firing rates of a network with stimulus-based connectivity, exposed to Gaussian mixtures (top row) and skewed Gaussians (bottom row). Firing rates are averaged over 400 sniff cycles at (a) $w = 0$, (b) $w = 2$, (c) $w = 10$ and (d) $w = 50$.

Figures A.24 shows average sniff cycles in networks with stimulus-based inhibition. These are similar to that of single-connected networks in that emergent rhythms are led by each stimulus' primacy set. However, all-to-all coupling allows rhythms to persist throughout the sniff, albeit with progressively less crispness. Figure A.25 shows the corresponding results of random and optimal discriminability at $T = 2$ and $T = 50$.

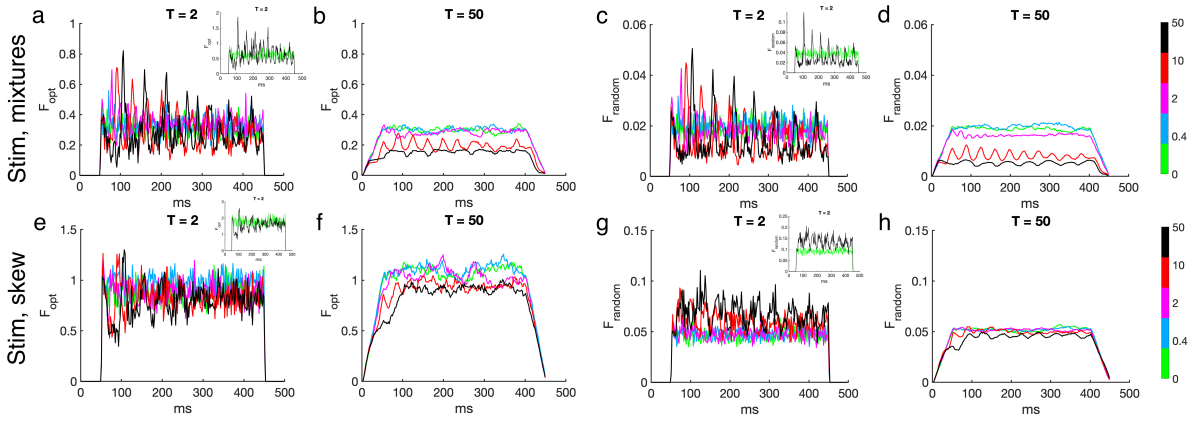


Figure A.25. $\mathcal{F}_{opt}$ and $\mathcal{F}_{random}$ for stimulus-based connected networks responding to Gaussian mixtures (top row) and skewed Gaussians (bottom row). Insets in (aceg) show data for only $w = 0$ and $w = 50$. Data is scaled with $T$.

APPENDIX B

# Exploratory Analyses

This appendix includes analyses that are not directly relevant to the main results of this project but may provide useful references for future work.

## B.1. Measurement error of Fisher information scales with measurement bin in Poisson networks.

In many of our analyses, we find that the variance of rates is slightly higher for small window sizes, especially for low or zero levels of inhibition (Figure 2.5). It is unclear whether this small uptick is a result of measurement error or is an inherent property of the system (e.g. that small sampling sizes intuitively pick up more variability from noisy spiking data, or that resetting voltage after a spikes produces an inherent refractory period). While we cannot determine an explicit analytical formula for the error in variance estimates from our network, we can analytically show that for a Poisson network with fixed mean, $\mathcal{F}_{opt}$ is overestimated at small measurement bins by a margin that decreases with increasing size of the measurement bin. These analyses provide an example of a scenario in which small deviations at small measurement bins may be a result of over- or under-estimation, rather than properties of the network.

Recall that we often compare network performance to that of Poisson spike trains of the same mean rates. In the following calculation we show that the deviation of the average estimated spike count $\hat{\mu}_i$ in response to stimulus $i$ from the true average $\mu_i$ leads

to an overestimation in Fisher information that decreases linearly with the size of the measurement bin.

Since Poisson neurons are assumed to be uncorrelated, the Fisher information in a network of $N$ neurons is the sum of contributions from each neuron. We assess the expected overestimation of a single neuron:

$$
\begin{aligned}
(\text{B.1}) \qquad E\left[\frac{(\hat{\mu}_1-\hat{\mu}_2)^2}{\hat{\mu}_1+\hat{\mu}_2}-\frac{(\mu_1-\mu_2)^2}{\mu_1+\mu_2}\right] &= E\left[\frac{(\hat{\mu}_1-\hat{\mu}_2)^2}{\hat{\mu}_1+\hat{\mu}_2}\right]-\frac{(\mu_1-\mu_2)^2}{\mu_1+\mu_2} \\
&= E\left[\frac{(\hat{\mu}_1+\hat{\mu}_2)^2}{\hat{\mu}_1+\hat{\mu}_2}\right]-E\left(\frac{4\hat{\mu}_1\hat{\mu}_2}{\hat{\mu}_1+\hat{\mu}_2}\right)-\frac{(\mu_1-\mu_2)^2}{\mu_1+\mu_2} \\
&= \mu_1+\mu_2-E\left(\frac{4\hat{\mu}_1\hat{\mu}_2}{\hat{\mu}_1+\hat{\mu}_2}\right)-\frac{(\mu_1+\mu_2)^2-4\mu_1\mu_2}{\mu_1+\mu_2} \\
&= -4E\left(\frac{\hat{\mu}_1\hat{\mu}_2}{\hat{\mu}_1+\hat{\mu}_2}\right)+\frac{4\mu_1\mu_2}{\mu_1+\mu_2} \\
&= 4\left(\frac{\mu_1\mu_2}{\mu_1+\mu_2}-E\left(\frac{\hat{\mu}_1\hat{\mu}_2}{\hat{\mu}_1+\hat{\mu}_2}\right)\right).
\end{aligned}
$$

The standard error of the mean is $\sqrt{\frac{\lambda}{n}}$ , where $\lambda$ is the Poisson parameter (in this case the expected number of spikes) and $n$ is the sample size. By the central limit theorem, $\hat{\mu}_i$ is normally distributed with standard deviation $\sqrt{\frac{\mu_i}{n}}$. We can thus write expected numbers of counts as the sums of the true counts and zero-mean error terms.

$$
(\text{B.2}) \qquad \hat{\mu}_1=\mu_1+\frac{\delta_1}{\sqrt{n}}, \qquad \hat{\mu}_2=\mu_2+\frac{\delta_2}{\sqrt{n}}.
$$

Then using straightforward algebraic calculations, the right hand side of the previous calculation is

$$
\begin{aligned}
(B.3) \qquad & 4\left(\frac{\mu_1\mu_2}{\mu_1+\mu_2} - E\left(\frac{\hat{\mu}_1\hat{\mu}_2}{\hat{\mu}_1+\hat{\mu}_2}\right)\right) \\
&= 4\left(\frac{\mu_1\mu_2}{\mu_1+\mu_2} - E\left(\frac{\left(\mu_1+\delta_1/\sqrt{n}\right)\left(\mu_2+\delta_2/\sqrt{n}\right)}{\left(\mu_1+\delta_1/\sqrt{n}\right)+\left(\mu_2+\delta_2/\sqrt{n}\right)}\right)\right) \\
&= 4\left(\frac{\mu_1\mu_2}{\mu_1+\mu_2} - E\left(A+B+C\right)\right),
\end{aligned}
$$

where

$$
(B.4) \qquad A = \frac{\mu_1\mu_2/(\mu_1+\mu_2)}{1+(\delta_1+\delta_2)/\left[(\mu_1+\mu_2)\sqrt{n}\right]}
$$

$$
(B.5) \qquad B = \frac{\mu_2\delta_1/\sqrt{n}+\mu_1\delta_2/\sqrt{n}}{\mu_1+\mu_2+\delta_1/\sqrt{n}+\delta_2/\sqrt{n}}
$$

$$
(B.6) \qquad C = \frac{\delta_1\delta_2/n}{\mu_1+\mu_2+\delta_1/\sqrt{n}+\delta_2/\sqrt{n}}
$$

Next, invoke the Taylor approximation

$$
(B.7) \qquad \frac{1}{1+\epsilon} \sim 1-\epsilon+\epsilon^2-O\left(\epsilon^3\right),
$$

using $\epsilon = \frac{\delta_1+\delta_2}{(\mu_1+\mu_2)\sqrt{n}}$.

The first terms inside the expectation simplifies as

$$E\left(\frac{\mu_1\mu_2/(\mu_1+\mu_2)}{1+(\delta_1+\delta_2)/\left[(\mu_1+\mu_2)\sqrt{n}\right]}\right) \tag{B.8}$$

$$= E\left(\frac{\mu_1\mu_2}{\mu_1+\mu_2}\left[1-\frac{\delta_1+\delta_2}{(\mu_1+\mu_2)\sqrt{n}}+\frac{(\delta_1+\delta_2)^2}{n(\mu_1+\mu_2)^2}-\text{h.o.t.}\right]\right) \tag{B.9}$$

$$= \frac{\mu_1\mu_2}{\mu_1+\mu_2} - E\left(\frac{\mu_1\mu_2}{\mu_1+\mu_2}\left(\frac{(\delta_1+\delta_2)^2}{n(\mu_1+\mu_2)^2}-\text{h.o.t.}\right)\right) \tag{B.10}$$

Second term:

$$E\left(\frac{\mu_2\delta_1/\sqrt{n}+\mu_1\delta_2/\sqrt{n}}{\mu_1+\mu_2+\delta_1/\sqrt{n}+\delta_2/\sqrt{n}}\right) = E\left(\left(\frac{\mu_2\delta_1+\mu_1\delta_2}{\sqrt{n}}\right)\left(\frac{1}{1+\frac{\delta_1+\delta_2}{(\mu_1+\mu_2)\sqrt{n}}}\right)\right) \tag{B.11}$$

$$= E\left(\left(\frac{\mu_2\delta_1+\mu_1\delta_2}{\sqrt{n}}\right)\left(1-\frac{\delta_1+\delta_2}{(\mu_1+\mu_2)\sqrt{n}}-\text{h.o.t.}\right)\right) \tag{B.12}$$

Third term:

$$E\left(\frac{\delta_1\delta_2/n}{\mu_1+\mu_2+\delta_1/\sqrt{n}+\delta_2/\sqrt{n}}\right) = E\left(\left(\frac{\delta_1\delta_2}{n(\mu_1+\mu_2)}\right)\left(\frac{1}{1+\frac{\delta_1+\delta_2}{(\mu_1+\mu_2)\sqrt{n}})}\right)\right) \tag{B.13}$$

$$= \frac{1}{n(\mu_1+\mu_2)}E\left((\delta_1\delta_2)\left(1-\frac{\delta_1+\delta_2}{(\mu_1+\mu_2)\sqrt{n}}+\text{h.o.t.}\right)\right) \tag{B.14}$$

Before we put things together, we note the following:

First, when $n$ is large $\frac{1}{n\sqrt{n}} \ll \frac{1}{n} \ll \frac{1}{\sqrt{n}}$.

Second, $E\left(\frac{\mu_2\delta_1+\mu_1\delta_2}{\sqrt{n}}\right) = 0$ due to linearity of expectation and the assumption that $\delta_1$ and $\delta_2$ are zero-mean Gaussian random variables.

Third, $E(\delta_1\delta_2) = 0$. To see this, note that $\delta_1 = \sqrt{n}(\hat{\mu}_1-\mu_1)$, $\delta_2 = \sqrt{n}(\hat{\mu}_2-\mu_2)$, so $E(\delta_1\delta_2) = nE[(\hat{\mu}_1-\mu_1)(\hat{\mu}_2-\mu_2)] = 0$, the covariance of independent random variables.

Fourth, $\frac{\delta_1}{\sqrt{n}} \sim Norm(0, \frac{\sqrt{\mu_1}}{\sqrt{n}})$, so $\delta_1 \sim Norm\left(0, \sqrt{\mu_1}\right)$ and $E[\delta_i^2] = \mu_i$.

The error term simplifies to

$$\frac{4}{n}\left\{\frac{2\mu_1\mu_2}{\mu_1+\mu_2}+\frac{\mu_1\mu_2\left(\mu_1+\mu_2\right)}{\left(\mu_1+\mu_2\right)^3}\right\}+\text{h.o.t..} \tag{B.15}$$

In summary, Fisher information in a Poisson network is overestimated by an amount that scales linearly with the expected number of spikes and by extension, the window size.

## B.2. Random linear separability $\mathcal{F}_{random}$ is invariant to translations of the projection line.

Consider a two-class discrimination problem between two noisy activity clouds in $N$-dimensional space. Consider a vector $\vec{c}_{rand}$ and consider the vector parallel to $\vec{c}_{rand}$ that goes through the centroid $[m_1, m_2, ..., m_N]$. This corresponding vector $[x_1, ..., x_N]^T$ is parametrized as

$$[x_1, ..., x_N]^T = [m_1, ..., m_N]^T + t\vec{c}_{rand}$$

and the resulting $\mathcal{F}_{random}$ is

$$\mathcal{F}_{random} = \frac{\left([x_1, ..., x_N]^T \cdot \Delta\vec{\mu}\right)^2}{[x_1, ..., x_N]^T \left(\sum_1 + \sum_2\right)[x_1, ..., x_N]}. \tag{B.16}$$

We can decompose the denominator as a sum over the eigenmodes of $\sum_1 + \sum_2$, indexed by $k$, where eigenvectors $\vec{e}_k$ are assumed to be normalized.

$$\mathcal{F}_{random} = \frac{\left([x_1, ..., x_N]^T \cdot \Delta\vec{\mu}\right)^2}{\sum_k \lambda_k \left([x_1, ..., x_N]^T \cdot \vec{e}_k\right)^2}. \tag{B.17}$$

The quantity $\left([x_1, ..., x_N]^T \cdot \Delta\vec{\mu}\right)$ is the signed magnitude of the projection of $\Delta\vec{\mu}$ onto $[x_1, ..., x_N]^T$. Similarly, the quantity $\left([x_1, ..., x_N]^T \cdot \vec{e}_k\right)$ is the magnitude of the projection

of $\vec{e}_k$ onto $[x_1, ..., x_N]^T$. By properties of projections, and since $[x_1, ..., x_N]^T \parallel \vec{c}_{rand}$, we have

$$\left([x_1, ..., x_N]^T \cdot \Delta\vec{\mu}\right) = \left(\vec{c}_{rand} \cdot \Delta\vec{\mu}\right) \tag{B.18}$$

$$\left([x_1, ..., x_N]^T \cdot \vec{e}_k\right) = \left(\vec{c}_{rand} \cdot \vec{e}_k\right), \tag{B.19}$$

and we conclude that

$$\mathcal{F}_{random} = \frac{\left([x_1, ..., x_N]^T \cdot \Delta\vec{\mu}\right)^2}{[x_1, ..., x_N]^T \left(\sum_1 + \sum_2\right) [x_1, ..., x_N]} = \frac{\left(\vec{c}_{rand} \cdot \left(\langle\vec{M}_1\rangle - \langle\vec{M}_2\rangle\right)\right)^2}{\vec{c}_{rand}^T \left(\sum_1 + \sum_2\right) \vec{c}_{rand}}. \tag{B.20}$$

## B.3. The benefit of induced noise correlations scales inversely to overall stimulus discriminability.

We consider how noise correlation impacts discriminability in networks whose principal neurons can be either excitatory or inhibitory. The purpose of this exploration was to assess the impact of correlations induced by lateral connectivities in general, beyond the inhibitory networks known in the olfactory system. Excitatory MCs represent disinhibition, for instance caused by MCs exciting cortical cells that inhibit GCs.

The network receives a symmetric pair of block stimuli. Each MC either delivers excitation or inhibition to all other cells in its own stimulus group, and either excitation or inhibition to all cells receiving the other stimulus. The coupling strength "within-block" is independently tuned from the coupling strength "across-block." For notational convenience, we often refer to inhibition within the same block as self-inhibition and inhibition to the other block as lateral-inhibition.

### B.3.1. Self-inhibition induces beneficial covariances even as discriminability is reduced overall.

In this first regime, mitral cells inhibit other cells in their same stimulus group and also inhibit other cells in the other stimulus group (Figure B.1a). Expectedly, the average difference in response between the two stimuli increases as lateral inhibition increases, and decreases as self-inhibition increases (Figure B.1b). Since one group of cells receives higher stimulation than the other, cells in that group tend to suppress the activity of the other group, enhancing the differences between them. Simultaneously, increasing the strength of self-inhibition causes cells to indirectly inhibit their own activities, and differences between average firing rates is diminished.

In (Figure B.1c), we hold the amount of lateral inhibition constant while increasing the amount of self inhibition (red to black lines). In doing so, $\mathcal{F}_{opt}$ and $\mathcal{F}_{uncorr}$ both decrease with the strength of self inhibition, and seemingly monotonically with window size. However, the proportion $\mathcal{F}_{opt}/\mathcal{F}_{uncorr}$ increases as we increase self-inhibition, suggesting that even though discriminability is dominated by the reduction in $\Delta\vec{\mu}$, inhibition induces pairwise covariances that are beneficial, especially at large coupling strengths.

A similar trend emerges when we swap out lateral inhibition for lateral disinhibition, whereby each MC inhibits the granule cells that inhibit MCs of the other stimulus group. In this scenario, $\Delta\vec{\mu}$ ends to decrease with self-inhibition, and also increases with the strength of lateral disinhibition (Figure B.2b, right to left) due to higher-excited mitral cells becoming disinhibited by weakly-stimulated MCs of the other group. As before, $\mathcal{F}_{opt}$ and $\mathcal{F}_{uncorr}$ trend downward with increasing self inhibition (Figure B.2c). Unlike in the previous scenario, discriminability trends upwards with increasing window size,

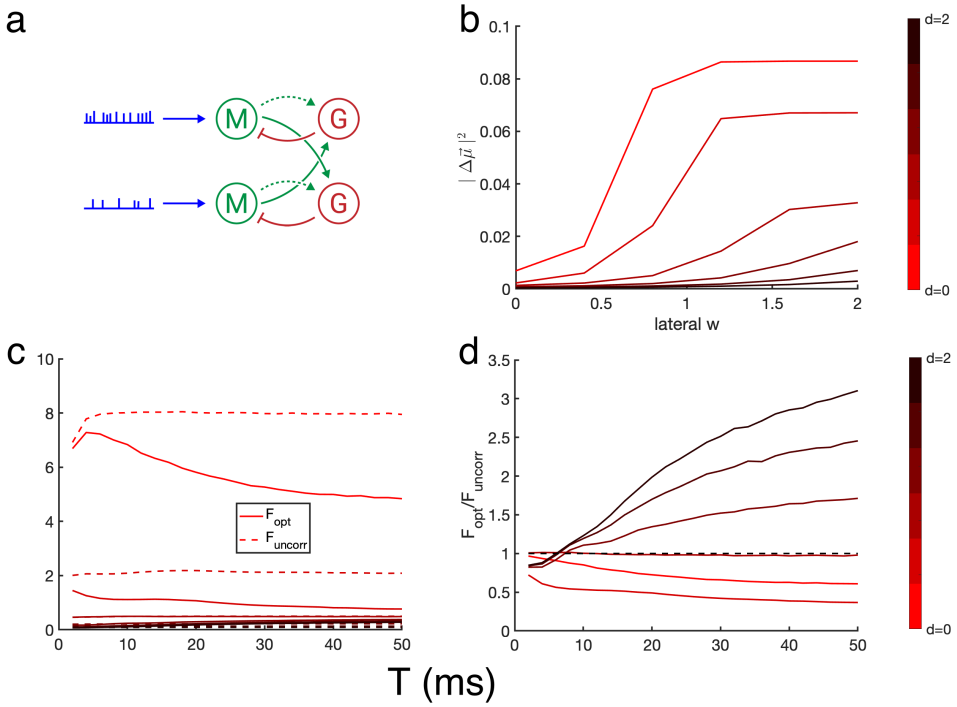


Figure B.1. (a) Network schematic, in which each M represents all MCs receiving a level of input. (b) Squared magnitude of $\Delta\vec{\mu}$. Horizontal axis is strength of lateral inhibition. Increasingly darker shades of red correspond to strength of self-inhibition. (c) $\mathcal{F}_{opt}$ and $\mathcal{F}_{uncorr}$ for measurement bins $T = 2$ to 50. Lateral inhibition is set to 1. (d) Proportion of $\mathcal{F}_{opt}$ to $\mathcal{F}_{uncorr}$ for measurement bins $T = 2$ to 50.

corresponding to covariances also becoming increasingly beneficial with larger measurement bin (Figure B.2d). Future work may address the nonmonotonic behavior of covariances within this regime, shown in Figure B.2d, which suggests the possibility of network rhythms emerging from patterns of network coupling.

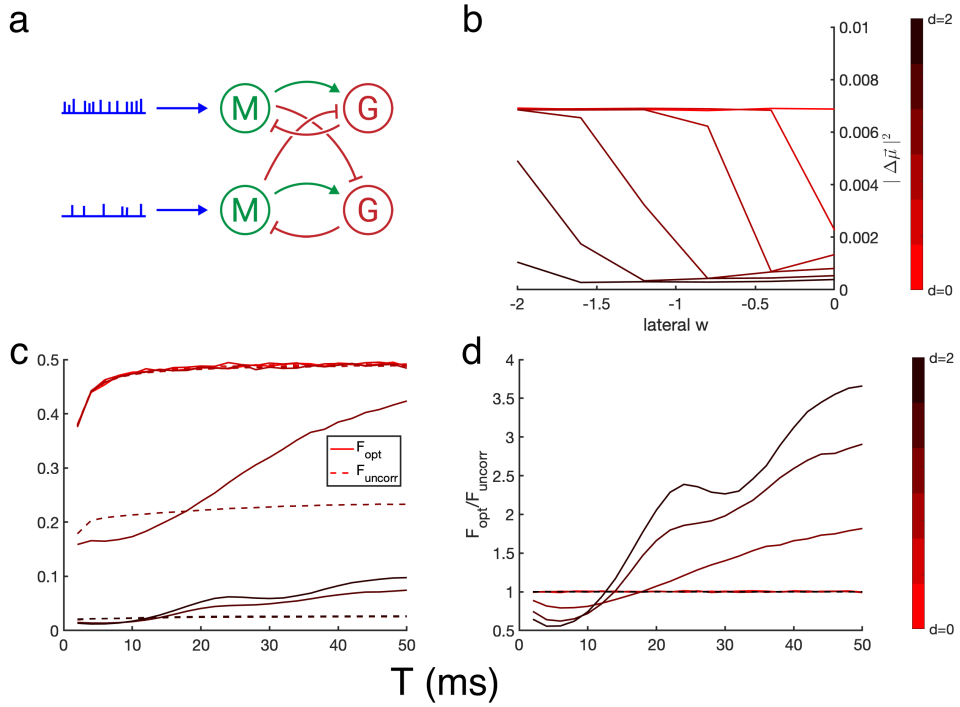


Figure B.2. (a) Network schematic, in which each M represents all MCs receiving a level of input. (b) Squared magnitude of $\Delta\vec{\mu}$. Horizontal axis is strength of lateral inhibition. Increasingly darker shades of red correspond to strength of self-inhibition. (c) $\mathcal{F}_{opt}$ and $\mathcal{F}_{uncorr}$ for measurement bins $T = 2$ to 50. Lateral inhibition is set to 1. (d) Proportion of $\mathcal{F}_{opt}$ to $\mathcal{F}_{uncorr}$ for measurement bins $T = 2$ to 50.

### B.3.2. Self disinhibition reduces discriminability and induces pairwise covariances that are not beneficial.

The final network contains self disinhibition paired with lateral excitation, which functionally resembles network interactions within the visual system. In this regime, mitral cells inhibit the granule cells that inhibit themselves, but excite the granule cells that inhibit MCs of the other group.

In Figure B.4, we plot the squared magnitude of $\Delta\vec{\mu}$ against the strength of lateral coupling, with darkening shades of blue corresponding to increasing amounts of self

disinhibition. The quantity $|\Delta\vec{\mu}|^2$ expectedly increases with lateral inhibition but unexpectedly trends lower with increasing self disinhibition. The latter result is unintuitive, since we'd expect that higher-stimulated MCs that indirectly disinhibit themselves would further bolster their activity above MCs of the weakly-stimulated group (Figure B.3). The reason for the downward trend is as follows. For nonzero lateral inhibition and zero self-inhibition (e.g. lateral $w = 1$, lightest blue line), highly-excited MCs reach a maximum activity level, capping out the amount of improvement they can hold over MCs in the other stimulus group. This maximum level is set by these highly-excited MCs fully suppressing weakly-stimulated MCs, whose corresponding GCs then provide no inhibition to the first group of cells.

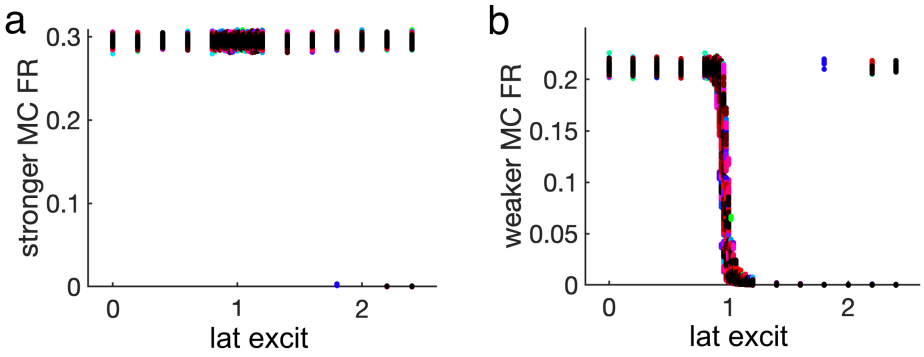


Figure B.3. Firing rates of MCs in a network with self-disinhibition. MCs inhibit their own GCs with $w_{mg} = -1.2$ (arbitrary units). Horizontal axis is strength of excitatory coupling from MCs towards other group (arbitrary units). Each dot represents one MC in one simulation. (a) Firing rate of highly-excited MCs. (b) Firing rates of weakly-excited MCs.

As the amount of self-disinhibition is increased (blue to black, Figure B.4b), the group of weakly-excited MCs occasionally dominate. This uncommon event is enabled because OSNs inputs are noisy, and occasionally, MCs in the weakly-stimulated group receive enough stimulation to "win out" over the other group. This inverted steady-state occurs

more frequently with higher amounts of self-disinhibition, which overall drives down the average mean difference across all trials. To conclude, higher self-disinhibition creates a bistability within the system in which either set of cells may become the dominating group.

Both $\mathcal{F}_{opt}$ and $\mathcal{F}_{uncorr}$ trend downward with increasing self disinhibition, and the proportion $\mathcal{F}_{opt}/\mathcal{F}_{uncorr}$ is less than 1 for most coupling strengths. These results are consistent with similar studies of lateral connectivity in the visual system that suggest self-excitation paired with lateral inhibition reduces total available information and induces pairwise correlations that further weaken discriminability [**12, 13, 14, 15, 16**].

**Note**: We do not consider networks in which MCs disinhibit both themselves and cells of the other stimulus group. Since the GCs in our network do not receive any input aside from MC excitation, in this network, the GCs receive no input altogether and hence do not spike. This scenario is not particularly interesting.

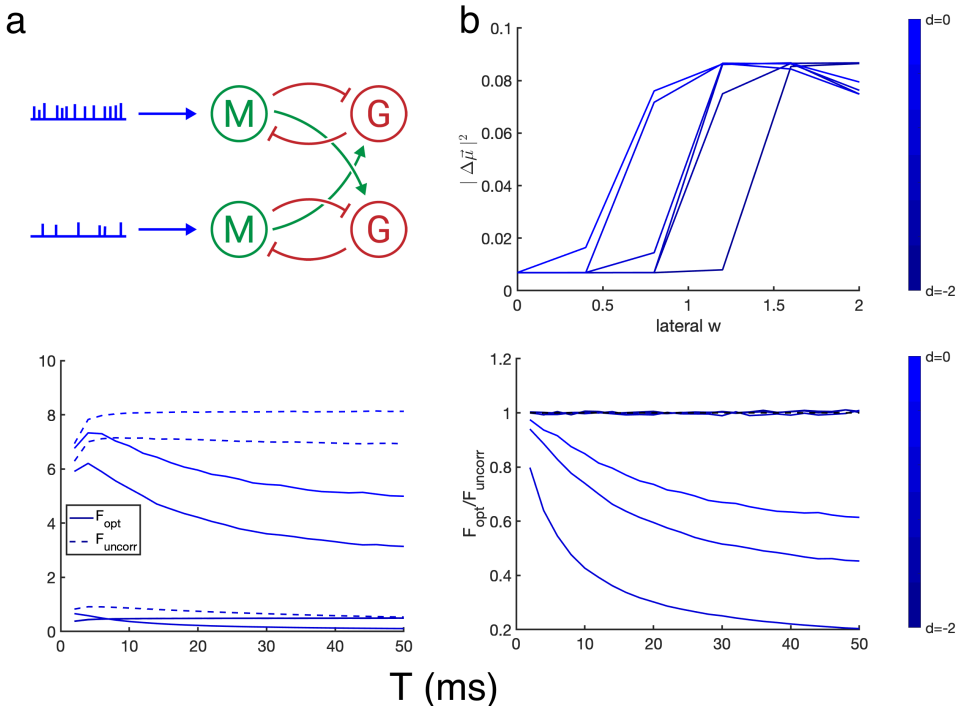


Figure B.4. (a) Network schematic, in which each M represents all MCs receiving a level of input. (b) Squared magnitude of $\Delta\vec{\mu}$. Horizontal axis is strength of lateral inhibition. Increasingly darker shades of blue correspond to strength of self-disinhibition. (c) $\mathcal{F}_{opt}$ and $\mathcal{F}_{uncorr}$ for measurement bins $T = 2$ to 50. Lateral inhibition is set to 1. (d) Proportion of $\mathcal{F}_{opt}$ to $\mathcal{F}_{uncorr}$ for measurement bins $T = 2$ to 50.

## B.4. Linear rate networks

We use a linear rate model to analytically understand how reciprocal and shared inhibition affect the separability of two similar stimuli, particularly through noise shaping.

### B.4.1. Variance of rates decreases slower than $(1+w)^{-2}$

Consider a simple linear rate equation that models the base interaction between a single granule cell $(G)$ and a single mitral cell $(M)$. MC activity decays exponentially, is reduced by granule cell activity via the coupling term $w$, and is incremented by stimulus $S$. GC activity is increased by mitral cell activity and changes in accordance to synaptic time constant $\tau$. Both the MC and GC receive independent zero-mean Gaussian noise, noted $\xi$ and $\zeta$ respectively.

$$\frac{dM}{dt} = -M - wG + S + \xi \tag{B.21}$$

$$\tau\frac{dG}{dt} = -G + M + \zeta. \tag{B.22}$$

We first assess how the difference in mean responses to two stimuli $(\Delta\mu)$ changes as a function of inhibitory coupling strength $w$. Denote $dW_1 = \xi dt$ and $dW_2 = \zeta dt$ and write the two differential equations in the form

$$dM = \left(-M + S - wG\right)dt + dW_1 \tag{B.23}$$

$$\tau dG = \left(-G + M\right)dt + dW_2. \tag{B.24}$$

Even if the stochastic differential system is not easily solvable, we can take the expectation of both sides. This system has the Martingale property, so taking the expectation

annihilates the stochastic terms, and the expected value of each variable satisfies the same differential equations equations as the original variables.

$$\frac{dE[M]}{dt} = -E[M] + S + w[G] \tag{B.25}$$

$$\tau\frac{dE[G]}{dt} = -E[G] + E[M]. \tag{B.26}$$

In steady state, the expected activities $E[M]$ and $E[G]$ satisfy

$$E[M] = S - wE[G] \tag{B.27}$$

$$E[G] = E[M]_+ \tag{B.28}$$

and we conclude

$$E[M] = \frac{S}{1+w}. \tag{B.29}$$

This implies that the mean difference activity in response to two stimuli decreases monotonically with $w$. That is,

$$\Delta\mu = \frac{S^{(1)} - S^{(2)}}{1+w}. \tag{B.30}$$

We next calculate how the covariance sum $\Sigma$ changes with inhibitory weight.

As we show (Appendix B.5.1), the variance of a signal can be calculated as a function of the power spectral density.

$$Var\left(\frac{1}{T}\int_0^T M(t)\ dt\right) = \frac{1}{\pi T^2}\int_{-\infty}^{\infty} |\ M(\omega)\ |^2\ \frac{1-\cos\omega T}{\omega^2}\ d\omega. \tag{B.31}$$

We endeavor to find the mitral cell's power spectral density $| M(\omega) |^2$.

$$\frac{d\hat{M}}{dt} = \frac{d}{dt}\int M(t)e^{-i\omega t}dt = -i\omega \int M(t)e^{-i\omega t}dt = i\omega\hat{M}. \tag{B.32}$$

The Fourier transform of (B.21) and (B.22) is

$$-i\omega\hat{M} = -\hat{M} - w\hat{G} + \hat{\xi} \tag{B.33}$$

$$-i\omega\tau\hat{G} = -\hat{G} + \hat{M} + \hat{\zeta} \tag{B.34}$$

$\hat{M}$ has the form

$$\hat{M} = \frac{-w\hat{\zeta} + \hat{\xi}\left(1 - i\omega\tau\right)}{\left(1 - i\omega\right)\left(1 - i\omega\tau\right) + w}$$

Taking their product, and denoting $\hat{\xi}\hat{\xi}^* = \hat{\xi}_0^2$ and $\hat{\zeta}\hat{\zeta}^* = \hat{\zeta}_0^2$, yields

$$| M(\omega) |^2 = E\left[\hat{M}\hat{M}^*\right] \tag{B.35}$$

$$= \frac{w^2\hat{\zeta}_0^2 + \hat{\xi}_0^2\left(1 + \omega^2\tau^2\right)}{\left(1 + w\right)^2 + \omega\left(\tau^2 + 1 - 2w\tau\right) + \omega^4\tau^2}. \tag{B.36}$$

In the second line, we use the property that the expectation of each noise term is 0 and that they are independent. Plugging this form into our equation for variance produces a formula for the variance of rates in terms of system parameters $w$, $\tau$, $\hat{\xi}_0^2$, and $\hat{\zeta}_0^2$,

(B.37)

$$Var\left(\frac{1}{T}\int_0^T M(t)\ dt\right) = \frac{1}{\pi T^2}\int_{-\infty}^{\infty}\left(\frac{w^2\hat{\zeta}_0^2 + \hat{\xi}_0^2\left(1 + \omega^2\tau^2\right)}{\left(1 + w\right)^2 + \omega\left(\tau^2 + 1 - 2w\tau\right) + \omega^4\tau^2}\right)\frac{1 - \cos\omega T}{\omega^2}\ d\omega.$$

Before computing the Fourier Transform of the system, we omitted the stimulus $S$ itself for computational convenience. In our analyses, $S$ is constant in time. Including this term produces the following expression for variance:

$$\text{(B.38)} \qquad Var\left(\frac{1}{T}\int_0^T M(t)\ dt\right)$$

$$\text{(B.39)} \qquad = \frac{1}{\pi T^2}\int_{-\infty}^{\infty}\left(\frac{w^2\hat{\zeta}_0^2 + \left(\hat{\xi}_0^2 + S^2\right)\left(1+\omega^2\tau^2\right)}{\left(1+w\right)^2 + \omega\left(\tau^2+1-2w\tau\right) + \omega^4\tau^2}\right)\frac{1-\cos\omega T}{\omega^2}\ d\omega - \left(\frac{S}{1+w}\right)^2.$$

In short, it adds a term that is proportional to $\frac{1}{(1+w)^2}$ and does not change how the integrand scales with $w$.

When $w = 0$, the power spectrum of the signal has a peak at low frequencies, whereas $w > 0$ shifts the peak to a higher frequency (Figure B.5). The variance, which is computed as a numerical integral over the PSD, consequently is lower for $w > 0$ (Figure B.6).

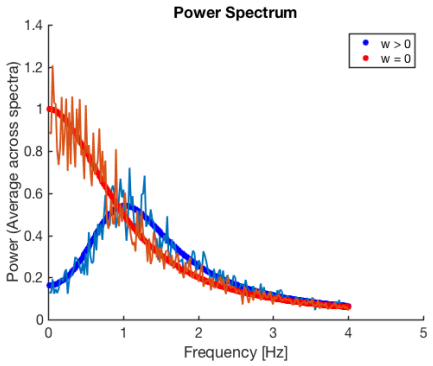


Figure B.5. As inhibition increases, power is reduced for each frequency. Figure is shown for $\hat{\zeta}_0^2 = 0, \hat{\xi}_0^2 = 1, \tau = 1$.

For $\hat{\zeta}_0^2 \neq 0$, the power spectral density is larger at higher frequencies and for larger inhibitory coupling $w$ (Figure B.7). In general, at larger inhibitory weights, the peak of the power spectrum occurs at higher frequency, and the peak is larger when there is

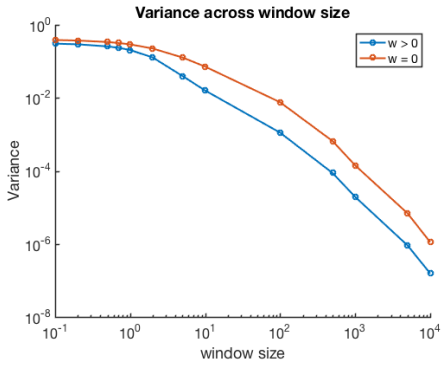


Figure B.6. Variance decreases with increasing window size. In the presence of inhibition, variance decreases more rapidly.

additional noise input to the granule cell. A slower GC (larger time constant) generates a power spectral density with peaks at smaller frequencies (Figure B.7).

For $\hat{\zeta}_0^2 = 0$ specifically, we plot the variance curves of the previous analysis with $(1+w)^2$ and found that the result increases with $w$, indicating that the original variance decreases with $w$ more slowly than $\frac{1}{(1+w)^2}$ (Figure B.8).

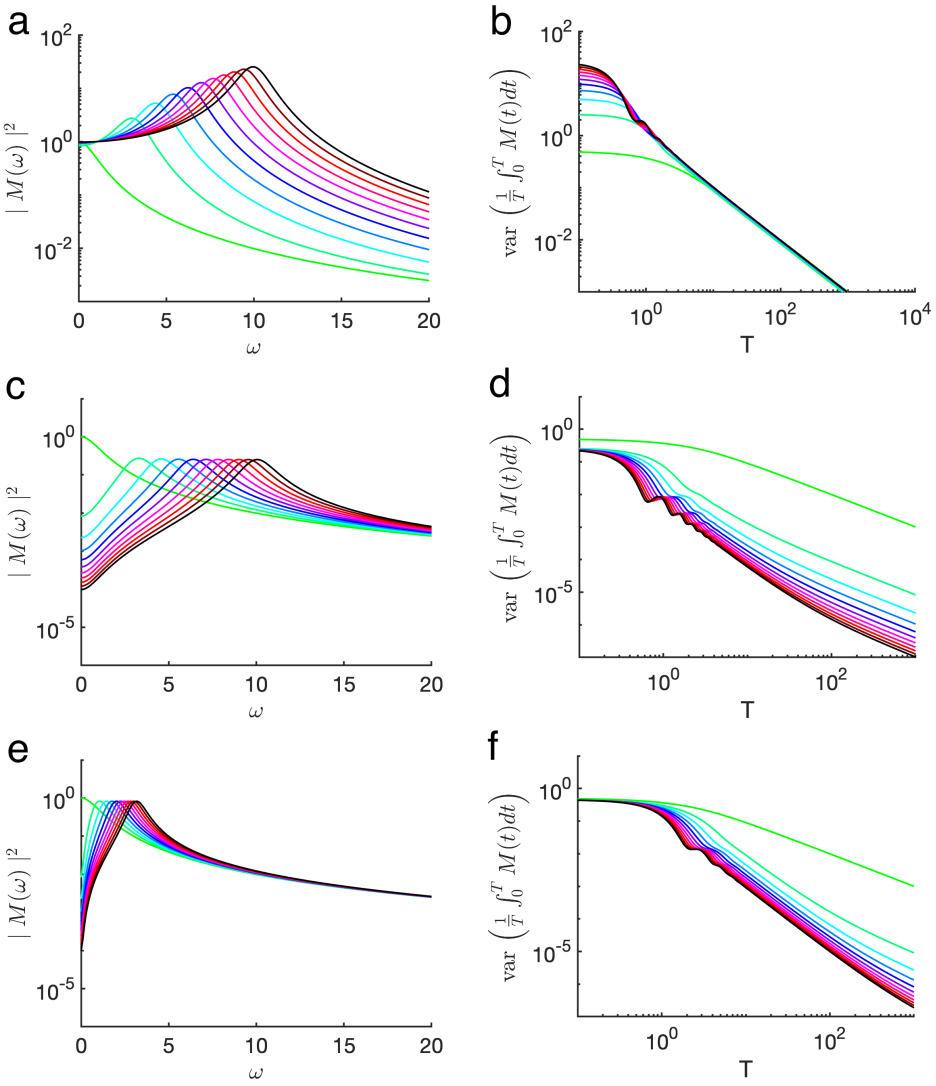


Figure B.7. Power spectral densities and corresponding variances of rates for (ab) $\hat{\zeta}_0^2 = 1$, $\tau = 1$, (cd) $\hat{\zeta}_0^2 = 0$, $\tau = 1$, and (ef) $\hat{\zeta}_0^2 = 0$, $\tau = 10$. Horizontal axis in first column is frequency in radians; horizontal axis in second column is $T$ in arbitrary units. Colored lines from green to black indicate increasing inhibitory weights from $w = 0$ to $w = 100$ (arbitrary units).

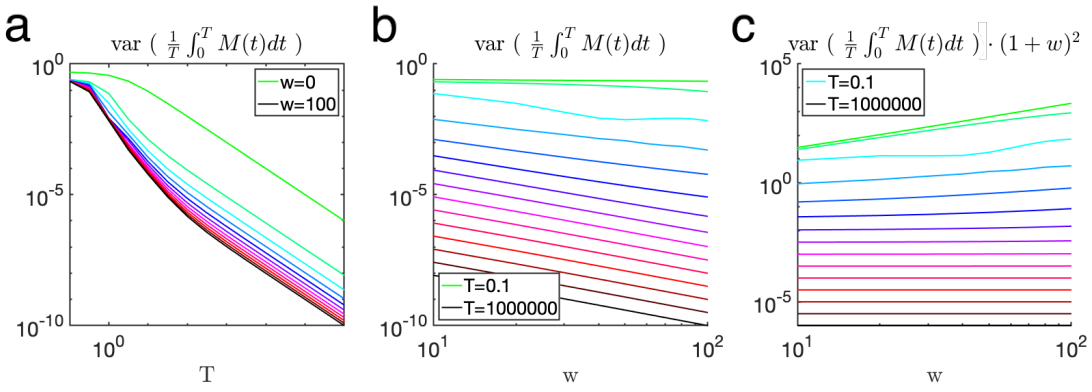


Figure B.8. (a) Variance of rates with $T$. Different color lines indicate increasing inhibitory coupling, from $w = 0$ to $w = 100$ (arbitrary units). (b) Variance of rates with $w$. Different color lines indicate increasing measurement bins from $T = 10^{-1}$ to $T = 10^{7}$ (artibrary units). (c) Same as (b) but multiplied by $(1 + w)^2$.

**B.4.1.1. Discriminability decays with $w$ in a single-connected network.** Consider a network of two mitral cells each reciprocally connected to a granule cell.

We expose the system to symmetric stimuli $S^{(1)}$ and $S^{(2)}$.

Mean difference in response monotonically decreases with $w$,

$$\Delta\vec{\mu} = \begin{pmatrix} E\left[M_1^{(1)}\right] - E\left[M_1^{(2)}\right] \\ E\left[M_2^{(1)}\right] - E\left[M_2^{(2)}\right] \end{pmatrix} = \frac{1}{1+w} \begin{pmatrix} S_1^{(1)} - S_1^{(2)} \\ S_2^{(1)} - S_2^{(2)} \end{pmatrix}. \tag{B.40}$$

$M_j^{(i)}$ is the response of mitral cell $j$ in response to stimulus $i$, with input $S_j^{(i)}$.

Since the cells do not share connections, the covariance matrix is a diagonal matrix whose entries are the variances of each MC, summed over both stimuli. We conclude that

$$\begin{aligned} \mathcal{F}_{opt} &\propto \frac{1}{(1+w)^2}\left[\frac{\left(S_1^{(1)} - S_1^{(2)}\right)^2}{Var\left(M_1\right)} + \frac{\left(S_2^{(1)} - S_2^{(2)}\right)^2}{Var\left(M_2\right)}\right] \\ &= \frac{2}{(1+w)^2}\left[\frac{\left(S_1^{(1)} - S_1^{(2)}\right)^2}{Var\left(M\right)}\right] \end{aligned} \tag{B.41}$$

by symmetry of the stimuli and because $Var\left(M_1\right) = Var\left(M_2\right)$. Note that since $Var(M)$ decreases more slowly than $\frac{1}{(1+w)^2}$ at small measurement bins, the overall effect is that $\mathcal{F}_{opt}$ decreases with increasing inhibitory weight $w$.

Simulations of the linear system show that $\mathcal{F}_{opt}$ indeed decreases with inhibitory weight for small measurement, and this effect is magnified for a slower granule cell response. Intuitively, a slower granule cell with a larger inhibitory weight induces rhythms in the mitral cell activity. A small measurement window that would otherwise only pick up

minimal variance content at low frequencies would then picks up greater variance, which reduces $\mathcal{F}_{opt}$. At sufficiently large window sizes, $\mathcal{F}_{opt}$ is unaffected by inhibitory weight.

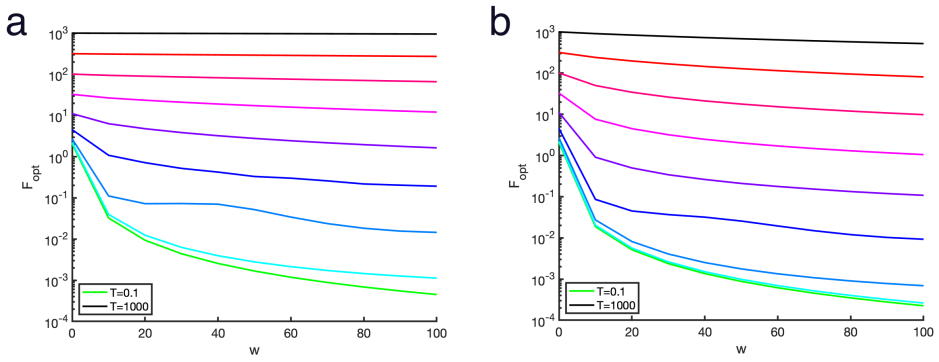


Figure B.9. Predicted $\mathcal{F}_{opt}$ for a 4-cell linear network connected with single reciprocal connections for a) $\tau = 1$ and b) $\tau = 10$. Colored lines from green to black indicate increasing $T$.

**B.4.1.2. $\Delta\vec{\mu}$ is independent of $w$ in a laterally-connected network.** We now slightly complicate our rate model to assess the effects of lateral inhibition on stimulus discriminability. Consider a system of two excitatory mitral cells that both excite the same granule cell and receive inhibition in turn. A common stimulus $S^{(i)}$ excites the two mitral cells at levels $S_1^{(i)}$ and $S_2^{(i)}$, and every neuron receives uncorrelated, zero-mean Gaussian noise.

$$\frac{dM_1}{dt} = -M_1 - wG + S_1^{(i)} + \xi_1 \tag{B.42}$$

$$\frac{dM_2}{dt} = -M_2 - wG + S_2^{(i)} + \xi_2 \tag{B.43}$$

$$\tau\frac{dG}{dt} = -G + (M_1 + M_2) + \zeta \tag{B.44}$$

We write $\Delta\mu$ in terms of system parameters. As before, we take the expectation of the system, which annihilates the stochastic terms,

$$0 = -E\left[M_1^{(i)}\right] - wE\left[G\right] + S_1^{(i)} \tag{B.45}$$

$$0 = -E\left[M_2^{(i)}\right] - wE\left[G\right] + S_2^{(i)} \tag{B.46}$$

$$0 = -E\left[G\right] + \left(E\left[M_1^{(i)}\right] + E\left[M_2^{(i)}\right]\right) \tag{B.47}$$

Straightforward algebra yields that for stimulus $i$,

$$E\left[M_2^{(i)}\right] = \frac{wS_2^{(i)} + S_2^{(i)} - wS_1^{(i)}}{1+2w} \tag{B.48}$$

$$E\left[M_1^{(i)}\right] = \frac{wS_1^{(i)} + S_1^{(i)} - wS_2^{(i)}}{1+2w} \tag{B.49}$$

which we use to expand and simplify $\Delta\mu$

$$\text{(B.50)} \qquad \Delta\mu = \begin{pmatrix} E\left[M_1^{(2)}\right] - E\left[M_1^{(1)}\right] \\ E\left[M_2^{(2)}\right] - E\left[M_1^{(2)}\right] \end{pmatrix}$$

$$= \begin{pmatrix} \frac{wS_1^{(2)}+S_1^{(2)}-wS_2^{(2)}}{1+2w} - \frac{wS_1^{(1)}+S_1^{(1)}-wS_2^{(1)}}{1+2w} \\ \frac{wS_2^{(2)}+S_2^{(2)}-wS_1^{(2)}}{1+2w} - \frac{wS_2^{(1)}+S_2^{(1)}-wS_1^{(1)}}{1+2w} \end{pmatrix}$$

$$= \frac{1}{1+2w} \begin{pmatrix} wS_1^{(2)} + S_1^{(2)} - wS_2^{(2)} - \left(wS_1^{(1)} + S_1^{(1)} - wS_2^{(1)}\right) \\ wS_2^{(2)} + S_2^{(2)} - wS_1^{(2)} - \left(wS_2^{(1)} + S_2^{(1)} - wS_1^{(1)}\right) \end{pmatrix}$$

$$= \begin{pmatrix} S_1^{(2)} - S_1^{(1)} \\ S_1^{(1)} - S_1^{(2)} \end{pmatrix}.$$

In the last line, we've made a symmetry assumption that $S_1^{(1)} = S_2^{(2)}$ and $S_1^{(2)} = S_1^{(1)}$. This is equivalent to supposing that the difference between our two stimuli is odd under reflections ($\Delta\vec{\mu}$ is odd under the exchange of mitral cells), rather than constant multiples of each other ($\Delta\vec{\mu}$ is even). (In these studies, we only consider stimuli that are symmetric.). We conclude that the mean response difference $\Delta\mu$ does not depend on $w$.

### B.4.2. Optimal linear discriminability is independent of $w$ in a laterally-connected network.

Shared inhibition induces pairwise correlations (and covariances) between the two mitral cells. We next calculate the pairwise covariance of rates between cells $M_1$ and $M_2$.

Without loss of generality, assume they each have an average value of zero.

(B.51)

$$
\begin{aligned}
&Cov\left(\frac{1}{T}\int_0^T M_1(t)dt, \frac{1}{T}\int_0^T M_2(t)dt\right) \\
&= E\left[\left(\frac{1}{T}\int_0^T M_1(t)dt\right)\left(\frac{1}{T}\int_0^T M_2(t)dt\right)\right] - \cancelto{0}{E\left[\frac{1}{T}\int_0^T M_1(t)dt\right]E\left[\frac{1}{T}\int_0^T M_2(t)dt\right]} \\
&= E\left[\frac{1}{T^2}\int_0^T\int_0^T M_1(t_1)M_2(t_2)dt_1dt_2\right] \\
&= \frac{1}{T^2}\int_0^T\int_0^T E\left[M_1(t_1)M_2(t_2)\right]dt_1dt_2.
\end{aligned}
$$

Under the assumption that $M_1$ and $M_2$ are stationary processes, $E\left[M_1(t_1)M_2(t_2)\right]$ is the cross-correlation between the two processes and depends only on the time difference $t_2 - t_1$, which we denote $\tau$.

(B.52)

$$
\frac{1}{T^2}\int_0^T\int_0^T E\left[M_1(t_1)M_2(t_2)\right]dt_1dt_2 = \frac{1}{T^2}\int_0^T\int_0^T E\left[\int_{-\infty}^{\infty} M_1\left(t'\right)M_2\left(t'+\tau\right)dt'\right]dt_1dt_2
$$

Isolate and simplify the expectation:

$$E\left[\int_{-\infty}^{\infty} M_1(t')\, M_2(t'+\tau)\, dt'\right] \tag{B.53}$$

$$= \int_{-\infty}^{\infty} E\left[\left(\frac{1}{2\pi}\int \hat{M}_1(\omega)e^{i\omega t'}d\omega\right)\left(\frac{1}{2\pi}\int \hat{M}_2(\omega')e^{i\omega' t'+i\omega'\tau}d\omega'\right)\right]dt'$$

Exchange the order of integration.

$$= \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int E\left[\hat{M}_1(\omega)e^{i\omega'\tau}\hat{M}_2(\omega')\left(\int e^{it'(\omega+\omega')}dt'\right)\right]d\omega' d\omega$$

$$= \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int E\left[\hat{M}_1(\omega)e^{i\omega'\tau}\hat{M}_2(\omega')\right](2\pi\delta(\omega+\omega'))\, d\omega' d\omega$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} E\left[\hat{M}_1(\omega)e^{i\omega'\tau}\hat{M}_2(-\omega)\right]d\omega$$

$$= \mathcal{F}^{-1}\left(E\left[\hat{M}_1\hat{M}_2^*\right]\right)(\tau)$$

We have thus so far written covariance in terms of the cross spectrum.

(B.54)

$$Cov\left(\frac{1}{T}\int_0^T M_1(t)dt, \frac{1}{T}\int_0^T M_2(t)dt\right) = \frac{1}{T^2}\int_0^T\int_0^T \mathcal{F}^{-1}\left(E\left[\hat{M}_1\hat{M}_2^*\right]\right)dt_1 dt_2$$

$$= \frac{1}{2\pi T^2}\int_0^T\int_0^T\left[\int E\left[\hat{M}_1\hat{M}_2^*\right]e^{i\omega(t_2-t_1)}d\omega\right]dt_1 dt_2 \tag{B.55}$$

Taking the Fourier transform of our original system yields

$$-i\omega\hat{M}_1 = -\hat{M}_1 - w\hat{G} + \hat{\xi}_1 \tag{B.56}$$

$$-i\omega\hat{M}_2 = -\hat{M}_2 - w\hat{G} + \hat{\xi}_2 \tag{B.57}$$

$$-i\omega\tau\hat{G} = -\hat{G} + \left(\hat{M}_1 + \hat{M}_2\right) + \hat{\zeta} \tag{B.58}$$

Isolate $\hat{M}_1$ and $\hat{M}_2^*$

$$\hat{G} = \frac{\left(\hat{\xi}_1 + \hat{\xi}_2\right) + \hat{\zeta}\left(1 - i\omega\right)}{\left(1 - i\omega\right)\left(1 - i\omega\tau\right) + 2w} \tag{B.59}$$

$$\hat{M}_1 = \frac{\hat{\xi}_1\left(\left(1 - i\omega\right)\left(1 - i\omega\tau\right) + w\right) - w\hat{\xi}_2 - w\hat{\zeta}\left(1 - i\omega\right)}{\left(1 - i\omega\right)\left[\left(1 - i\omega\right)\left(1 - i\omega\tau\right) + 2w\right]} \tag{B.60}$$

$$\hat{M}_2^* = \frac{-w\hat{\xi}_1^* + \hat{\xi}_2^*\left[\left(1 + i\omega\right)\left(1 + i\omega\tau\right) + w\right] - w\hat{\zeta}^*\left(1 + i\omega\right)}{\left(1 + i\omega\right)\left[\left(1 + i\omega\right)\left(1 + i\omega\tau\right) + 2w\right]} \tag{B.61}$$

In taking their product, we use the $\delta-$correlation properties of all the noise terms

$$\hat{M}_1\hat{M}_2^* \tag{B.62}$$

$$= \frac{-w\hat{\xi}_{1,0}\left(\left(1 - i\omega\right)\left(1 - i\omega\tau\right) + w\right) - w\hat{\xi}_{2,0}\left(\left(1 + i\omega\right)\left(1 + i\omega\tau\right) + w\right) + w^2\hat{\zeta}_0^2\left(1 + \omega^2\right)}{\left(1 + \omega^2\right)\left[\left(1 + \omega^2\right)\left(1 + \omega^2\tau^2\right) + 4w\left(1 - \omega^2\tau\right) + 4w^2\right]} \tag{B.63}$$

We now make an important observation about this quantity. The product $\hat{M}_1\hat{M}_2^*$ is complex, whereas the covariance of two real quantities must itself be real. Therefore, the inner integral $\int E\left[\hat{M}_1\hat{M}_2^*\right]e^{i\omega(t_2-t_1)}d\omega$ must also be real.

Note that the quantity $\hat{M}_1\left(\omega\right)\hat{M}_2^*\left(\omega\right)$ is equal to $\hat{M}_1^*\left(-\omega\right)\hat{M}_2\left(-\omega\right)$. Let $\omega' = -\omega$ and $d\omega' = -d\omega$.

$$\int_{-\infty}^{\infty} E\left[\hat{M}_1(\omega)\hat{M}_2^*(\omega)\right]e^{i\omega(t_2-t_1)}d\omega = \int_{-\infty}^{\infty} E\left[\hat{M}_1^*(-\omega)\hat{M}_2(-\omega)\right]e^{i\omega(t_2-t_1)}d\omega \tag{B.64}$$

$$= \int_{\infty}^{-\infty} E\left[\hat{M}_1^*(\omega')\hat{M}_2(\omega')\right]e^{-i\omega'(t_2-t_1)}\left(-d\omega'\right) \tag{B.65}$$

$$= \int_{-\infty}^{\infty} E\left[\hat{M}_1^*(\omega')\hat{M}_2(\omega')\right]e^{-i\omega'(t_2-t_1)}d\omega' \tag{B.66}$$

$$= \int_{-\infty}^{\infty} E\left[\hat{M}_1^*(\omega')\hat{M}_2(\omega')\right]e^{i\omega'(t_1-t_2)}d\omega', \tag{B.67}$$

which lets us conclude that taking the complex conjugate of the random variable $\hat{M}_1(\omega)\,\hat{M}_2^*(\omega)$ is equivalent to time-reversing the original quantity. We therefore have the two equivalent statements

$$E\left[\int_{-\infty}^{\infty} M_1(t')\,M_2(t'+\tau)\,dt'\right] = E\left[\int_{-\infty}^{\infty} M_1(t'+\tau')\,M_2(t')\,dt'\right] \tag{B.68}$$

$$\mathcal{F}^{-1}E\left[\hat{M}_1(\omega)\hat{M}^*{}_2(\omega)\right](t_2-t_1) = \mathcal{F}^{-1}E\left[\hat{M}_1^*(\omega')\hat{M}_2(\omega')\right](t_1-t_2)\,. \tag{B.69}$$

For convenience, we can replace $\mathcal{F}^{-1}\left(E\left[\hat{M}_1\hat{M}_2^*\right]\right)$ in our covariance computation with $\frac{1}{2}\mathcal{F}^{-1}\left(E\left[\hat{M}_1\hat{M}_2^* + \hat{M}_1\hat{M}_2^*\right]\right)$ without changing the value of the covariance but annihilating all the imaginary terms.

Simplifying,

$$Cov\left(\frac{1}{T}\int_0^T M_1(t)dt, \frac{1}{T}\int_0^T M_2(t)dt\right) \tag{B.70}$$

$$= \frac{1}{T^2}\int_0^T\int_0^T \frac{1}{2}\mathcal{F}^{-1}\left(E\left[\hat{M}_1\hat{M}_2^* + \hat{M}_1\hat{M}_2^*\right]\right)dt_1dt_2 \tag{B.71}$$

$$= \frac{1}{T^2}\int_0^T\int_0^T \frac{1}{2(2\pi)}\int I\times e^{i\omega(t_2-t_1)}d\omega dt_1dt_2. \tag{B.72}$$

,

where

$$I = \frac{-w\hat{\xi}_{1,0}\left(2-2\omega^2\tau+2w\right) - w\hat{\xi}_{2,0}\left(2-2\omega^2\tau+2w\right) + 2w^2\hat{\zeta}_0^2\left(1+\omega^2\right)}{\left(1+\omega^2\right)\left[\left(1+\omega^2\right)\left(1+\omega^2\tau^2\right)+4w\left(1-\omega^2\tau\right)+4w^2\right]} \tag{B.73}$$

Exchanging the order of integration and simplify as in the single cell case gives

$$Cov\left(\frac{1}{T}\int_0^T M_1(t)dt, \frac{1}{T}\int_0^T M_2(t)dt\right) \tag{B.74}$$

$$= \frac{1}{\pi T^2}\int I \times \left(\frac{1-\cos\omega T}{\omega^2}\right) d\omega, \tag{B.75}$$

This yields an analytical formula for the covariance of two MCs that share inhibition.

$\mathcal{F}_{opt}$ can be written

$$\mathcal{F}_{opt} = \left(\sum_j (\Delta\vec{\mu}\cdot\vec{e}_j)\,\vec{e}_j\right)^T \Sigma^{-1}\left(\sum_k (\Delta\vec{\mu}\cdot\vec{e}_k)\,\vec{e}_k\right) \tag{B.76}$$

$$= \left(\sum_j (\Delta\vec{\mu}\cdot\vec{e}_j)\,\vec{e}_j\right)^T \left(\sum_k \frac{1}{\lambda_k}(\Delta\vec{\mu}\cdot\vec{e}_k)\,\vec{e}_k\right) \tag{B.77}$$

$$= \sum_j \frac{1}{\lambda_j}(\Delta\vec{\mu}\cdot\vec{e}_j)^2\,. \tag{B.78}$$

Our stimuli are symmetric, so $\Delta\vec{\mu}$ is an odd vector. We further assume that the noise to each of the two mitral cells have the same variance, i.e. $\hat{\xi}_0^2 = \hat{\xi}_{1,0}^2 = \hat{\xi}_{2,0}^2$, and denote

$$|\,M(\omega)\,|^2 = |\,M_1(\omega)\,|^2 = |\,M_2(\omega)\,|^2\,. \tag{B.79}$$

Then the covariance of sums $\Sigma$ is a symmetric matrix

$$\begin{bmatrix} a & c \\ c & a \end{bmatrix} \tag{B.80}$$

with determinant $a^2 - c^2$ and eigenvalues $\lambda_1,\ \lambda_2 = a + c$ and $a - c$, corresponding to eigenvectors $\begin{bmatrix} 1 & 1 \end{bmatrix}^T$ and $\begin{bmatrix} 1 & -1 \end{bmatrix}^T$, respectively. For odd $\Delta\mu$, only the odd eigenvector with corresponding eigenvalue $\lambda_j = a - c$ contributes to $\mathcal{F}_{opt}$.

The diagonal term $a$ is the sum of variances of both cells; the offdiagonal term $c$ is the pairwise covariance. They can each be computed analytically from the integral formulation:

(B.81)

$$
\begin{aligned}
a &= \frac{1}{\pi T^2} \int_{-\infty}^{\infty} \mid M(\omega) \mid^2 \frac{1 - \cos \omega T}{\omega^2}\, d\omega \\
&= \frac{1}{\pi T^2} \int_{-\infty}^{\infty} \left( \frac{\hat{\xi}_0^2 \left( (1+\omega^2)(1+\omega^2\tau^2) + 2w(1-\omega^2\tau) + 2w^2 \right) + w^2 \hat{\zeta}_0^2 (1+\omega^2)}{(1+\omega^2)\left[ (1+\omega^2)(1+\omega^2\tau^2) + 4w(1-\omega^2\tau) + 4w^2 \right]} \right) \frac{1 - \cos \omega T}{\omega^2}\, d\omega \\
c &= \frac{1}{\pi T^2} \int_{-\infty}^{\infty} \left( \frac{-2w\hat{\xi}_0^2 (1 - \omega^2\tau + w) + w^2 \hat{\zeta}_0^2 (1+\omega^2)}{(1+\omega^2)\left[ (1+\omega^2)(1+\omega^2\tau^2) + 4w(1-\omega^2\tau) + 4w^2 \right]} \right) \left( \frac{1 - \cos \omega T}{\omega^2} \right) d\omega.
\end{aligned}
$$

We next take $\hat{\xi}_0^2 = \tau = 1$ and compute the odd eigenvalue $a - c$, the only eigenvalue that contributes to $\mathcal{F}_{opt}$. A straightforward integral computation shows that $\hat{\zeta}_0$ drops out of the calculation; the noise input to the GC does not impact discriminability. Furthermore, $a - c$ is independent of $w$

(B.82)
$$a - c = \frac{1}{\pi T^2} \int_{-\infty}^{\infty} \frac{1}{1+\omega^2} \left( \frac{1 - \cos \omega T}{\omega^2} \right) d\omega$$

(B.83)
$$= \frac{T + e^{-T} - 1}{T^2}.$$

In other words, the inhibitory coupling between the GC and the two MCs does not reduce noise in the direction parallel to $\Delta\vec{\mu}$, only in the direction orthogonal. $\mathcal{F}_{opt}$ is unaffected by inhibitory strength.

The other eigenvalue, $a + c$, corresponds to the even mode.

$$a + c = \frac{1}{\pi T^2} \int_{-\infty}^{\infty} \left( \frac{1 + 2w^2 \hat{\zeta}_0^2 + \omega^2}{4w^2 - 4w\left(-1 + \omega^2\right) + \left(1 + \omega^2\right)^2} \right) \left( \frac{1 - \cos \omega T}{\omega^2} \right) d\omega. \tag{B.84}$$

This quantity does not contribute to optimal linear discrimination.

To conclude,

$$\mathcal{F}_{opt} = \left( \frac{T + e^{-T} - 1}{T^2} \right)^{-1} \left[ \begin{pmatrix} S_1^{(2)} - S_1^{(1)} \\ S_1^1 - S_1^2 \end{pmatrix} \cdot [-1, 1]^T \right] \tag{B.85}$$

$$= \frac{T^2}{-1 + e^{-T} + T} \left[ 2 \left( S_1^{(2)} - S_1^{(1)} \right)^2 \right]. \tag{B.86}$$

We plot the values of $\frac{1}{2}\left(M_1^* M_2 + M_2 M_1^*\right)$, the pairwise covariance between two cells,variances of each neuron individually, and both eigenvalues for $\hat{\zeta}_0^2 = 1$ and $\hat{\zeta}_0^2 = 0$, respectively (Figure B.10). All plots use $\hat{\xi}_{1,0}^2 = \hat{\xi}_{2,0}^2 = 1$, $\tau = 1$. When there is noise input to the granule cells ($\hat{\zeta}_0^2 \neq 0$), both the variance of each cell individually and the pairwise covariance of the cells increases with inhibitory coupling. The eigenvalue corresponding to the even eigenvector also increases with $w$. When there is no noise input to the granule cells, i.e. $\hat{\zeta}_0^2 = 0$ and mitral cell activity is the only thing driving their evolution, both variance and covariance decrease with $w$, and the even eigenvalue follows suit. In both cases, the qualitative behavior of the covariance follows from the shape of the corresponding power

spectral densities. Specifically, at small window sizes the integral computation of the covariance picks up frequency content at both small and large frequencies, whereas larger window sizes pick up only the frequency content at small frequencies. For cross-spectra that have a peak, increasing the window size beyond the frequency of the peak correlates with a sudden drop in covariance. We note in particular that when $\hat{\zeta}_0^2 \neq 0$, increasing inhibitory couplings creates a larger, higher peak in the PSD.

Notably, the odd eigenvalue decreases with window size but does not change with $w$, regardless of whether the granule cell receives external noise.

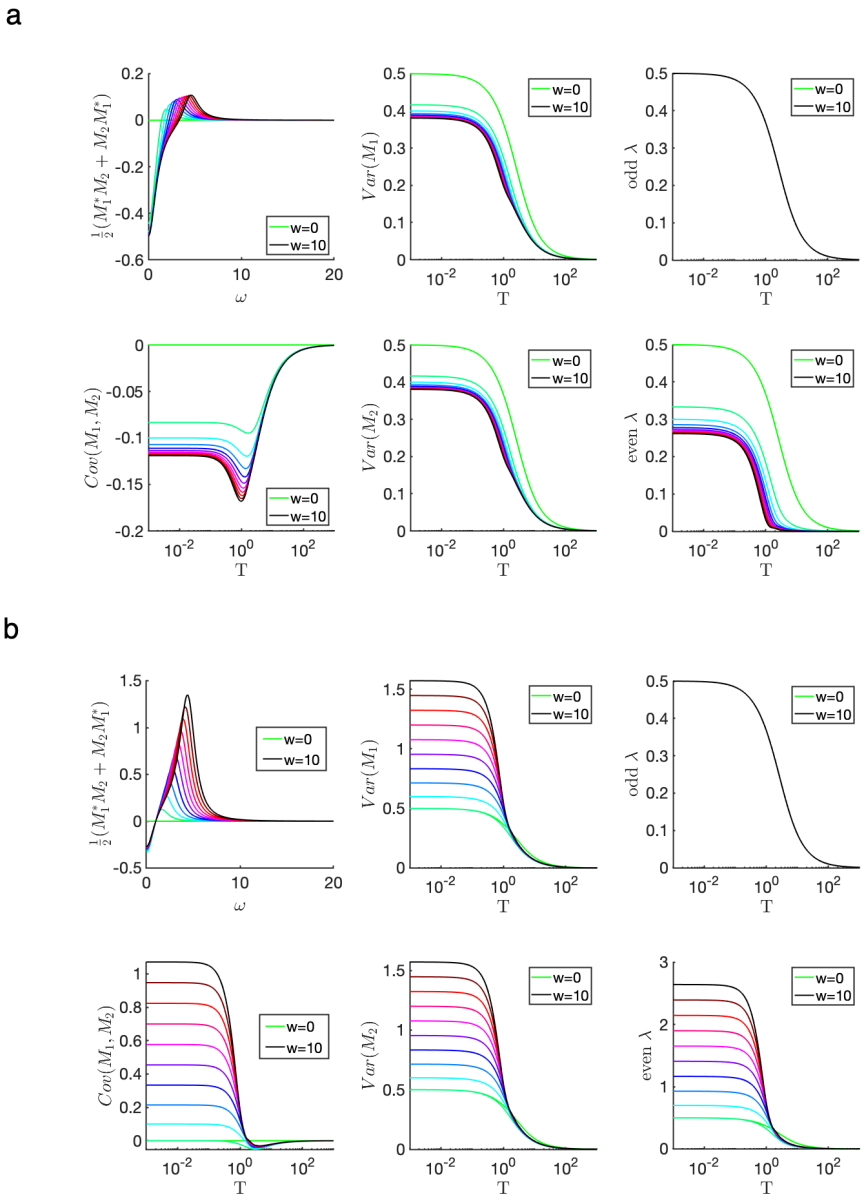


Figure B.10. Elements of the covariance matrix for a 2-MC linear system. (a) $\hat{\zeta}_0^2 = 0$. Top left: $\frac{1}{2}(M_1^* M_2 + M_2 M_1^*)$. Bottom left: Pairwise covariance of the two cells. Top middle: variance of mitral cell 1 for various $T$. Bottom middle: variance of mitral cell 2. Top right: Eigenvalue corresponding to the odd mode of the covariance matrix. Bottom right: Eigenvalue corresponding to the even mode of the covariance matrix. (b) Same as for (a) but for $\hat{\zeta}_0^2 = 1$. In all figures, colored lines from light green to black indicate increasing inhibitory weights from $w = 0$ to $w = 10$.

A few observations:

For $w = 0$ (light green line), pairwise covariance is zero regardless of whether there is noise input to the granule cells. This is consistent with the lack of common influence on the two cells, and they function as two independently driven neurons with no covariance.

When the granule cells do not receive noise input (i.e. $\hat{\zeta}_0^2 = 0$), covariance is negative, indicating that the two mitral cells are anticorrelated. When there is noise in the granule cells, covariance at small window sizes shows a positive correlation. Covariance at larger window sizes shows a negative correlation. In all our simulations, we suppose $\hat{\zeta}_0^2 = 0$.

In summary, we have the prediction that for single connected neurons,

(B.87) $$| \Delta\vec{\mu} | \propto \frac{1}{1+w}$$

(B.88) $Var(M)$ decreases slower than $\frac{1}{(1+w)^2}$, especially at small $T(FigureB.8)$.

(B.89) $$\mathcal{F}_{opt} = \frac{2}{(1+w)^2}\left[\frac{(S_{2,1} - S_{1,1})^2}{Var\,(M)}\right].$$

$\mathcal{F}_{opt}$ decreases with $w$ at small window sizes and is barely affected by $w$ at large measurement windows. When the granule cell evolves with a larger time constant $\tau$, it takes larger measurement windows yet before $\mathcal{F}_{opt}$ is unaffected by $w$. For any fixed inhibition, $\mathcal{F}_{opt}$ increases with measurement bin.

For globally connected neurons,

(B.90) $$F_{opt} = \frac{T^2}{-1 + e^{-T} + T}\left[2\,(S_{2,1} - S_{1,1})^2\right],$$

which monotonically increases with window size $T$ and does not change with inhibitory coupling $w$.

### B.4.3. Linear rate model simulations

We numerically evolve a rate model according to the equations described above.

First, we simulate dynamics of a single-connected network with two mitral cells and two granule cells connected reciprocally. System parameters are $\hat{\xi}_0^2 = 1$, $\hat{\zeta}_0^2 = 0$, and $\tau = 1$, and inhibitory weights are taken from $w = 0$ to $w = 50$ (arbitrary units). After each time series is produced, we take a sufficiently large number of samples of a fixed measurement bin, from which we compute $\mathcal{F}_{opt}$ as well as fraction of samples correctly classified. Chosen measurement bins range from $T = 5$ ms to $T = 200$ ms.

In globally-connected networks, we use the same model parameters as in single-connected networks. In order to keep the same amount of inhibition comparable to that of the single connection case, we scale the weight matrix in each direction by the total number of recipient cells. Indeed, we find that neither $\mathcal{F}_{opt}$ nor fraction correctly classified changes with inhibitory weight. Even though the average value of mitral cells decreases with inhibition, their difference is preserved. Furthermore, we verify that the dominant odd of the covariance matrix decreases monotonically with window size and is unaffected by inhibition. Taken together, we conclude that global inhibition preserves Fisher linear information even as overall variance content is reduced.

We briefly comment on the impacts in the limit of an infinitely fast granule cell response ($\tau \to 0$),

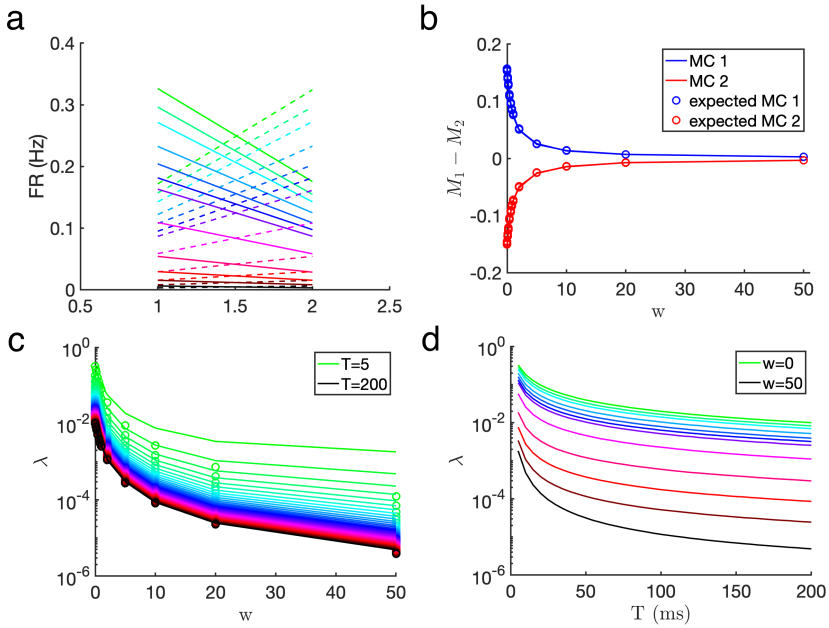


Figure B.11. Single-connected linear rate network. (a) Firing rates for mitral cells 1 and 2 for stimulus 1 (solid lines) and 2 (dashed lines). Colored lines from green to black indicate increasing inhibitory weight from $w = 0$ to $w = 50$ (arbitrary units). (b) Difference in average responses for different inhibitions. Solid lines are the results of numerical computations; open circles indicate analytical predictions. (c) Eigenvalue of the odd eigenvector with inhibitory weight. Green to black lines indicate increasing window sizes. Open circles indicate analytical predictions (d) Eigenvalue of the odd eigenvector with window size $T$. Green to black lines indicate increasing inhibitory strengths from $w = 0$ to $w = 50$.

The two governing systems of equations for single and global connections, respectively, becomes

$$\frac{dM}{dt} = -(1+w)M + S + \xi, \qquad \frac{dM_1}{dt} = -(1+w)M_1 - wM_2 + S_1^{(i)} + \xi_1 \tag{B.91}$$

$$\frac{dM_2}{dt} = -(1+w)M_2 - wM_1 + S_2^{(i)} + \xi_2. \tag{B.92}$$

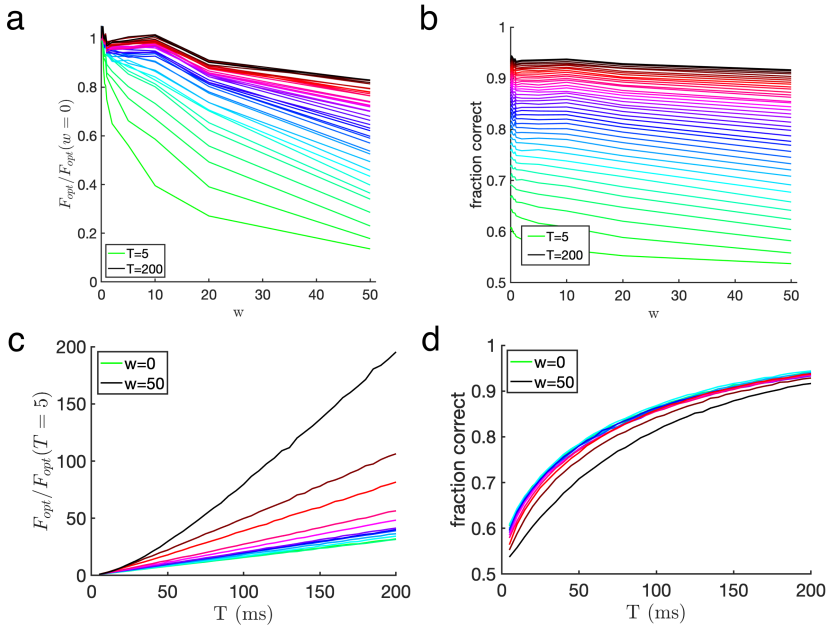


Figure B.12. Discriminability and classification performance of a single-connected linear rate network. (a) Proportional change of $\mathcal{F}_{opt}$ relative to that of a network with $w = 0$. Colored lines from green to black indicate increasing $T$. (b) Fraction correctly classified using linear discriminant analysis across inhibitory weights. Colored lines from green to black indicate increasing measurement bin. (c) $\mathcal{F}_{opt}$ for different measurement bins $T$. Colored lines from green to black indicate increasing inhibitory weight, from $w = 0$ to $w = 50$ (arbitrary units). (d) Fraction correctly classified across measurement bins. Colored lines from green to black indicate increasing inhibitory weight, from $w = 0$ to $w = 50$.

For the single-connected network, $\mathcal{F}_{opt}$ does not obviously decrease with inhibition as before, although the rate at which it improves with increasing $T$ is larger for stronger inhibition $w$. Both $\mathcal{F}_{opt}$ and fraction correctly classified increase with $T$.

In the global network, discriminability is still preserved across different inhibitory coupling $w$ and increases monotonically with $T$.

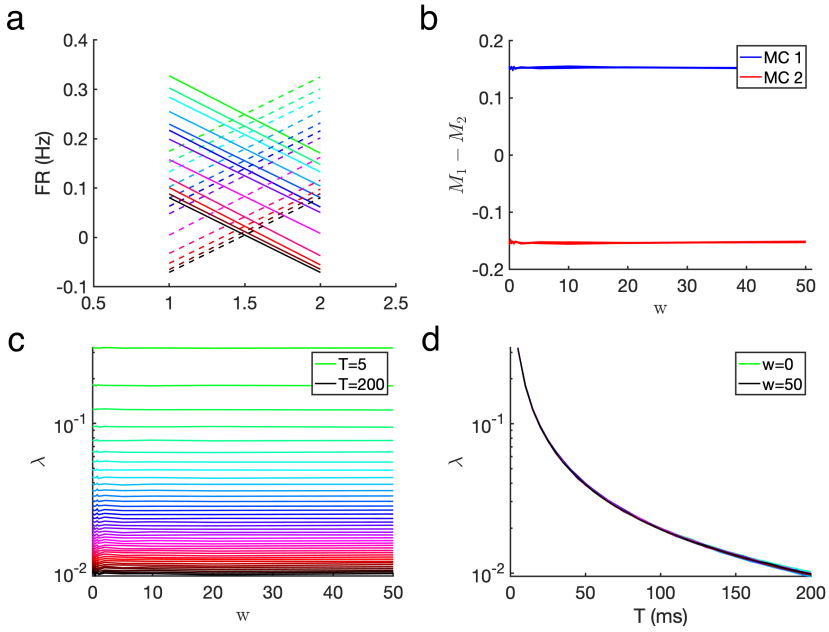


Figure B.13. Same as Figure B.11 but for globally-connected networks

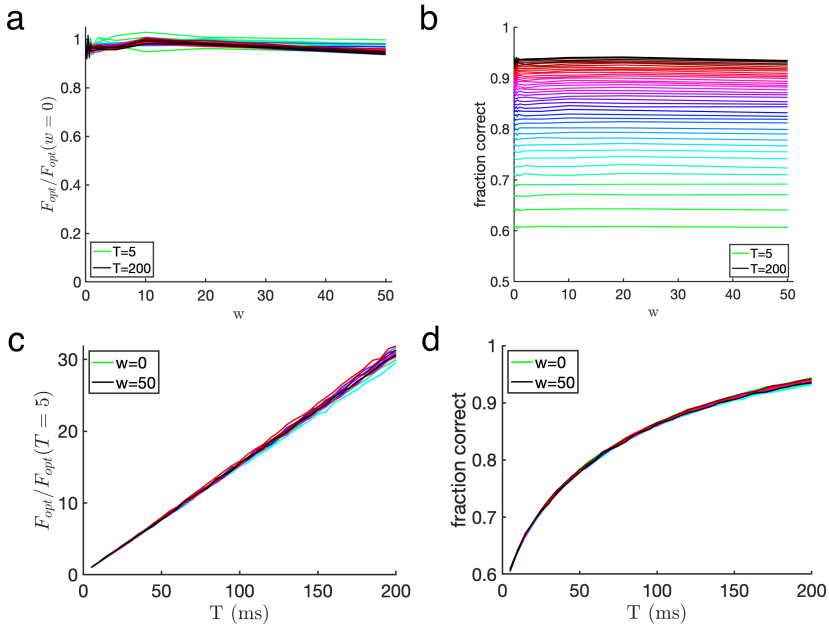


Figure B.14. Same as Figure B.12 but for globally-connected networks.

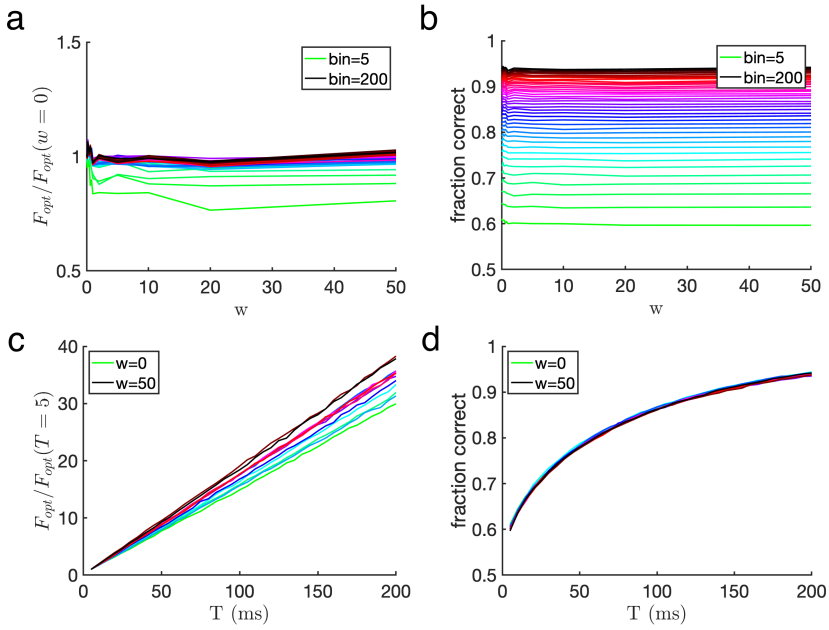


Figure B.15. Same as for Figure B.12 but with infinitely-fast granule cell.

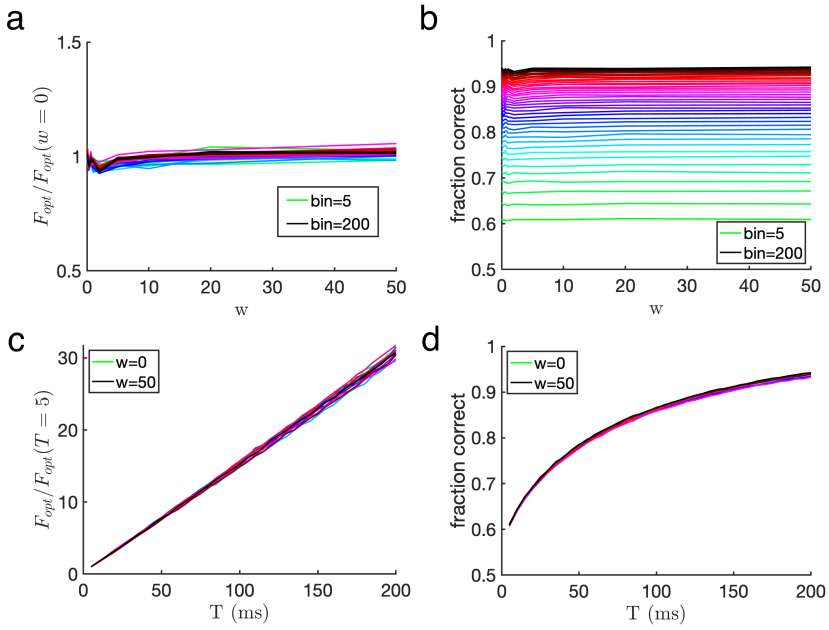


Figure B.16. Same as for Figure B.14 but with infinitely-fast granule cell.

### B.4.4. Inhibition improves $\mathcal{F}_{random}$ but not $\mathcal{F}_{opt}$ in globally connected linear networks.

One key feature of $\mathcal{F}_{random}$ is that it reflects changes in how noise correlations are shaped, not just along the optimal direction. In a globally connected network, $w$ reduces noise in the direction orthogonal to $\Delta\vec{\mu}$ (Figure B.17). This translates into an improvement in $\mathcal{F}_{random}$ (Figure B.18), which reflects the noise reduction along that direction. As we increase the number of MCs in the system however, this noise reduction has less of an effect. Put another way, $\mathcal{F}_{random}$ improves because the variability in a single eigenmode is decreased; at larger systems, this effect is barely noticeable (Figure B.19).

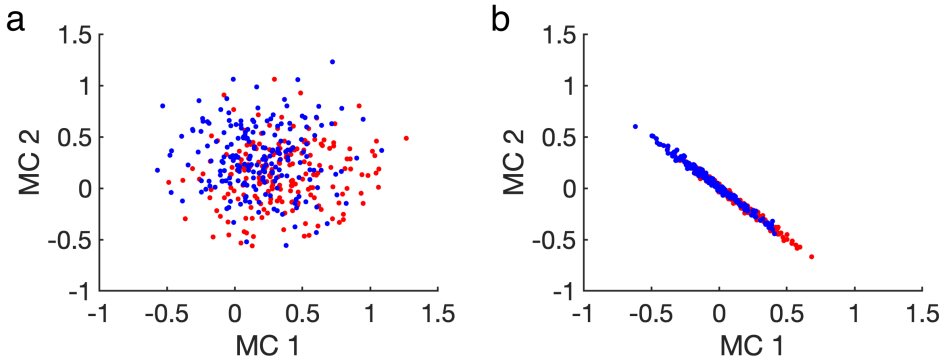


Figure B.17. (a) Scatterplot of sample rates measured in 200ms windows with no inhibition (b) global inhibition

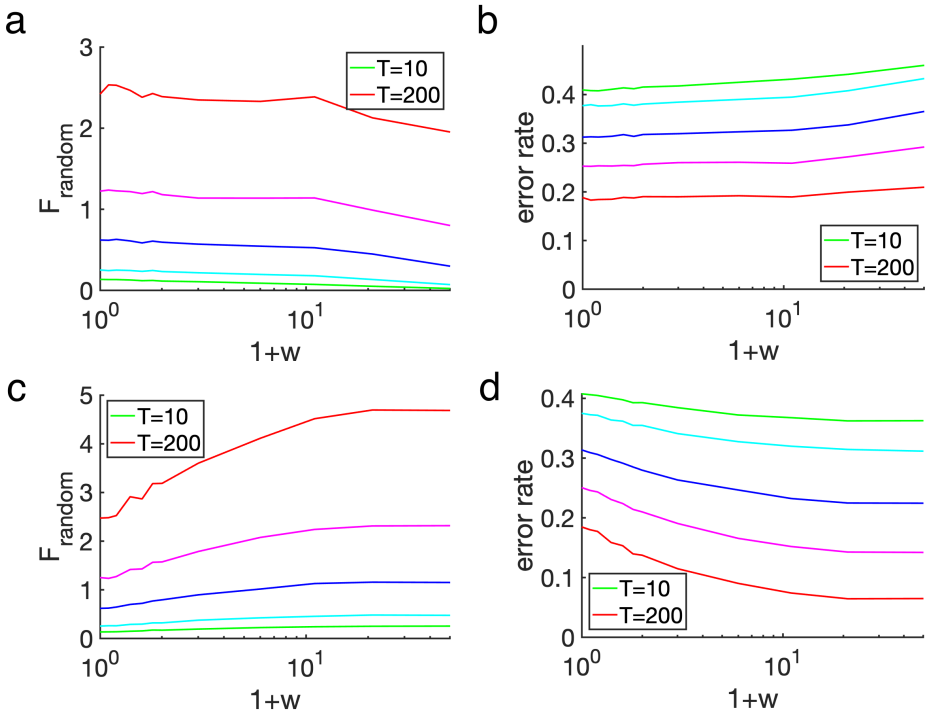


Figure B.18. Random discriminability in 2-MC network (a) Mean $\mathcal{F}_{random}$ for single-connected network. Different colored lines indicate increasing window size, from $T$ = 10 to $T$ = 200. (b) Corresponding error rate for each measurement bin in (a). (cd) same as (ab) but for globally-connected network.

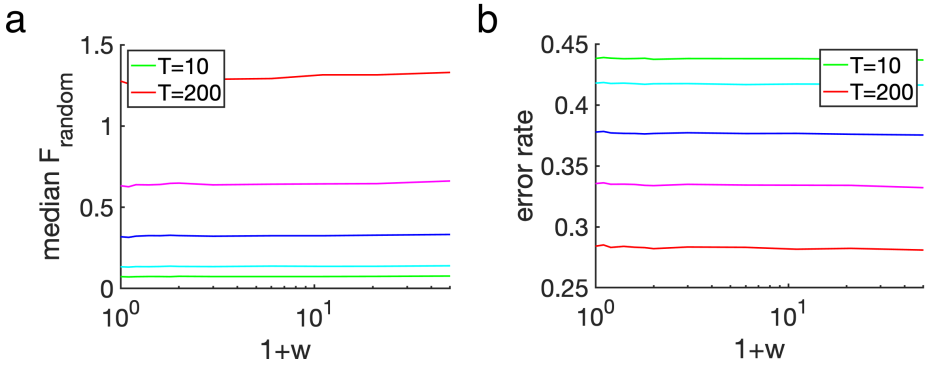


Figure B.19. Same as in Figure B.18cd but for 50-MC globally-connected linear network.

## B.5. Power spectral density derivations

### B.5.1. Power spectral density can be calculated from the variance

We derive a formula for the variance of a signal as a function of the power spectral density:

$$Var\left(\frac{1}{T}\int_0^T M(t)\ dt\right) = \frac{1}{\pi T^2}\int_{-\infty}^{\infty} \mid M(\omega) \mid^2 \frac{1-\cos\omega T}{\omega^2}\ d\omega. \tag{B.93}$$

Let $M_0(t)$ be a time series for which the integral over a finite interval of size $T$ gives the number of spikes. $M_0(t)$ can either be a continuous variable representing the time-averaged firing rate of a neuron or a discrete sum of Dirac delta functions. Let $M(t)$ be the result of subtracting its mean.

Then the variance of firing rates measured in windows of size $T$ can be written

$$\begin{aligned} Var\left(\frac{1}{T}\int_0^T M(t)\ dt\right) &= \frac{1}{T^2} Var\left(\int_0^T M(t)\ dt\right) \\ &= \frac{1}{T^2}\left\{ E\left[\left(\int_0^T M(t)\ dt\right)^2\right] - \underbrace{E\left[\left(\int_0^T M(t)\ dt\right)\right]^2}_{=0} \right\} \\ &= \frac{1}{T^2} E\left[\left(\int_0^T M(t)\ dt\right)^2\right] \\ &= \frac{1}{T^2} E\left[\left(\int_0^T M(t)\ dt\right)\left(\int_0^T M(t')\ dt'\right)\right] \\ &= \frac{1}{T^2}\int_0^T\int_0^T \underbrace{E\left[M(t)M(t')\right]}_{=A(t-t')}\ dtdt' \\ &= \frac{1}{T^2}\int_0^T\int_0^T A(t-t')\ dtdt', \end{aligned} \tag{B.94}$$

where $A(t-t')$ is the autocorrelation function.

We next invoke the Wiener-Kinchin theorem, which states that the autocorrelation of a time series is equal to the inverse Fourier transform of the power spectrum.

$$\begin{aligned}
\text{(B.95)}\quad \frac{1}{T^2}\int_0^T\int_0^T A(t-t')\,dtdt' &= \frac{1}{T^2}\int_0^T\int_0^T \frac{1}{2\pi}\int_{-\infty}^{\infty} |\,M(\omega)\,|^2\, e^{i\omega(t-t')}\,d\omega dtdt' \\
&= \frac{1}{2\pi T^2}\int_{-\infty}^{\infty} |\,M(\omega)\,|^2 \left(\int_0^T\int_0^T e^{i\omega(t-t')}\,dtdt\right)' d\omega \\
&= \frac{1}{2\pi T^2}\int_{-\infty}^{\infty} |\,M(\omega)\,|^2 \left(\frac{2}{\omega^2} - \frac{2}{\omega^2}\underbrace{\left(\frac{e^{i\omega T}-e^{-i\omega T}}{2}\right)}_{=\cos\omega T}\right) d\omega \\
&= \frac{1}{\pi T^2}\int_{-\infty}^{\infty} |\,M(\omega)\,|^2\,\frac{1-\cos\omega T}{\omega^2}\,d\omega.
\end{aligned}$$

Note that

$$\text{(B.96)}\qquad \int_{-\infty}^{\infty}\frac{1}{\pi T}\left(\frac{1-\cos\omega T}{\omega^2}\right)d\omega = 1$$

and the integrand biases towards lower frequencies for higher values of $T$ (Figure B.20). Specifically, this means that larger window sizes pick up more variance at lower frequencies of the power spectrum, and that variance decreases faster when there is a greater volume of activity at higher $\omega$.

### B.5.2. Power spectral density indicates amount of frequency content.

Consider a discrete time signal $x(t)$ of length $N$. Assuming it is periodic, we can write it as the composition of periodic functions of different finite frequencies.

$$\text{(B.97)}\qquad x_i = x(t_i) = \bar{x} + \sum_{n=1}^{N}\left(a_n\cos(\omega_n t_i) + b_n\sin(\omega_n t_i)\right)$$

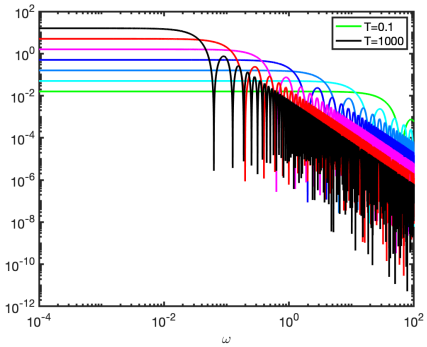


Figure B.20. The term $\left(\frac{1-\cos\omega T}{\omega^2}\right)$ for different frequencies. Colored lines from light green to black indicate exponentially increasing bin sizes $T$.

where$\bar{x}$ is the mean value of the signal, $\omega_n$ is the $n$th frequency, and $a_n$ and $b_n$ control the amplitudes of each mode.

By definition, the variance of $x(t_i)$ is the expected squared difference of the signal from its mean.

$$
\begin{aligned}
E[(x_i-\bar{x})^2] &= \frac{1}{N-1}\sum_{i=1}^{N}\left(\sum_{n=1}^{N} a_n\cos(\omega_n t_i)+b_n\sin(\omega_n t_i)\right)^2 \\
&= \frac{1}{N-1}\sum_{i=1}^{N}\sum_{n=1}^{N}\left(a_n\cos(\omega_n t_i)+b_n\sin(\omega_n t_i)\right)^2 \text{ by orthogonality} \\
&= \frac{1}{N-1}\sum_{n=1}^{N}\sum_{i=1}^{N}\left(a_n\cos(\omega_n t_i)+b_n\sin(\omega_n t_i)\right)^2 \\
&= \frac{1}{N-1}\sum_{n=1}^{N}\frac{N}{2}\left(a_n^2+b_n^2\right) \\
&= \frac{N}{2(N-1)}\sum_{n=1}^{N}\left(a_n^2+b_n^2\right),
\end{aligned}
\tag{B.98}
$$

where the second line follows by orthogonality and the fourth line uses known trigonometric properties.

The term $(a_n^2 + b_n^2)$ is by definition the power spectral density of $x(t)$ , denoted $\mid \hat{x}(\omega) \mid^2$. Each term in the sum indicates the contribution to the total variability for that particular frequency. It represents the amount of each frequency content present in the signal.

## B.6. Adaptive exponential integrate and fire neuron with refractory period

An earlier iteration of this project used an adaptive integrate and fire neuron with absolute refractory period, similar to that of [**32**]. It is similar to the model described in the main body of the thesis, except with two additional terms, shown below in red.

After a spike, neuron voltage was held at resting potential for 5 ms.

$$C\frac{dv}{dt} = -g_L\left(v - E_L\right) + g_L\Delta_T e^{\frac{v-v_T}{\Delta_T}} - w_{ad} + z + I \tag{B.99}$$

$$v > V_T \rightarrow v = E_L \text{ for 5 ms,} \tag{B.100}$$

It includes a hyperpolarizing adaptation current $w_{ad}$ whose evolution equation is

$$\tau_{w_{ad}}\frac{d}{dt}w_{ad} = a(v - E_L) - w_{ad}, \tag{B.101}$$

where $a$ is a parameter and $\tau_{w_{ad}}$ is the time constant of the adaptation of the neuron. Upon a spike, $w_{ad}$ is incremented by amount $b$.

The other addition is $z$, which is a depolarizing spike afterpotential. Upon a spike, $z$ is set to a value $I_{sp}$ and decays otherwise with time constant $z$.

$$\tau_z\frac{d}{dt}z = -z. \tag{B.102}$$

The third and final adaptation mechanism is an adaptive threshold. $V_T$ is set to $V_{T_{max}}$ following a spike and decays to $V_{T_{rest}}$ with a time constant $\tau_{V_T}$. That is,

$$\tau_{V_T}\frac{d}{dt}V_T = -(V_T - V_{T_{rest}}) \tag{B.103}$$

This neuron's average firing rate saturates with increasing excitation, and its frequency-input curve is concave down. Correspondingly, inhibition always has a larger effect on neurons that are more weakly stimulated. As a result, any amount of inhibition increases their average differences and also increases $\mathcal{F}_{opt}$.

There is a connectivity analyzed here known in the literature as "Mexican hat," characterized by MCs disinhibiting its direct neighbors but inhibiting cells that are further away. This connectivity has the effect of sharpening the activity responses, and is implemented as

$$W(x) = A_e \exp\left(-\frac{x^2}{\sigma_e^2}\right) - A_i \exp\left(-\frac{x^2}{\sigma_i^2}\right), \tag{B.104}$$

where $x$ is a spatial coordinate and represents for instance, the lateral distance between granule and mitral cells. This is equivalent to the difference of two Gaussian humps.

In all networks regardless of connectivity, discriminability increases with measurement duration (Figures B.21 and B.22). In these scenarios, percent correct (PC) is near 100%.

One of the reasons we opted for a simpler model was that in these simulations, it was difficult to understand clearly what exactly made discriminability improve under some conditions, and not others. For instance, we intuitively hypothesized that global connectivity would improve the discriminability of skewed Gaussians because it extracted larger differences between MC responses (Figure B.22, top row, second column). However, small amounts of inhibition actually worsened $\mathcal{F}_{opt}$ but larger amounts improved it.

Also, while inhibition generally changed the eigenvalues of the covariance matrix $\Sigma$ (Figure B.23, Figure B.24), it was difficult to pinpoint how $\sum$ terms contributed to discriminability. For example, inhibition in most connectivity regimes increased the magnitude of

covariance eigenvalues in networks exposed to Gaussian mixtures, but since contributions scale with the inverse of eigenvalues, it is the smallest ones whose change has the greatest impact. Altogether, these data encouraged us to consider a simpler model with which we could more precisely analyze performance.

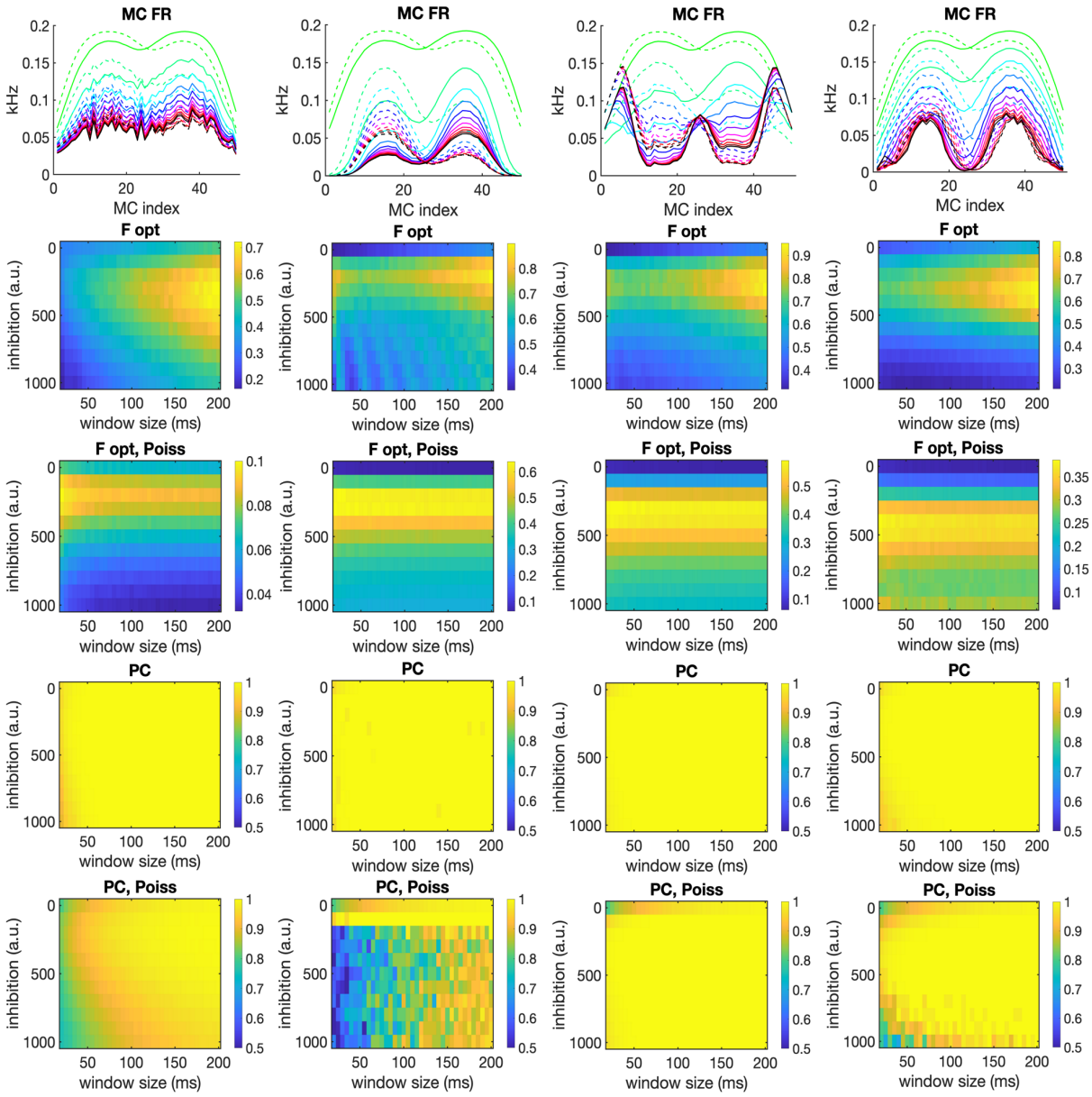


Figure B.21. Data from networks of AdExIF neurons exposed to Gaussian mixtures. Columns from left to right indicate single, global, stimulus-based, and Mexican hat connectivity, respectively. Rows correspond to average MC firing rate of stimulus 1 (solid line) and stimulus 2 (dashed line) at increasing inhibitions (green to black, arbitrary units), $\mathcal{F}_{opt}$ over measurement duration (horizontal axis) and inhibition (vertical axis), $\mathcal{F}_{opt}$ for the corresponding Poisson spike trains with same mean rate, percent correct using the optimal classifier on network data, percent correct using the optimal classifier on Poisson spike train data.

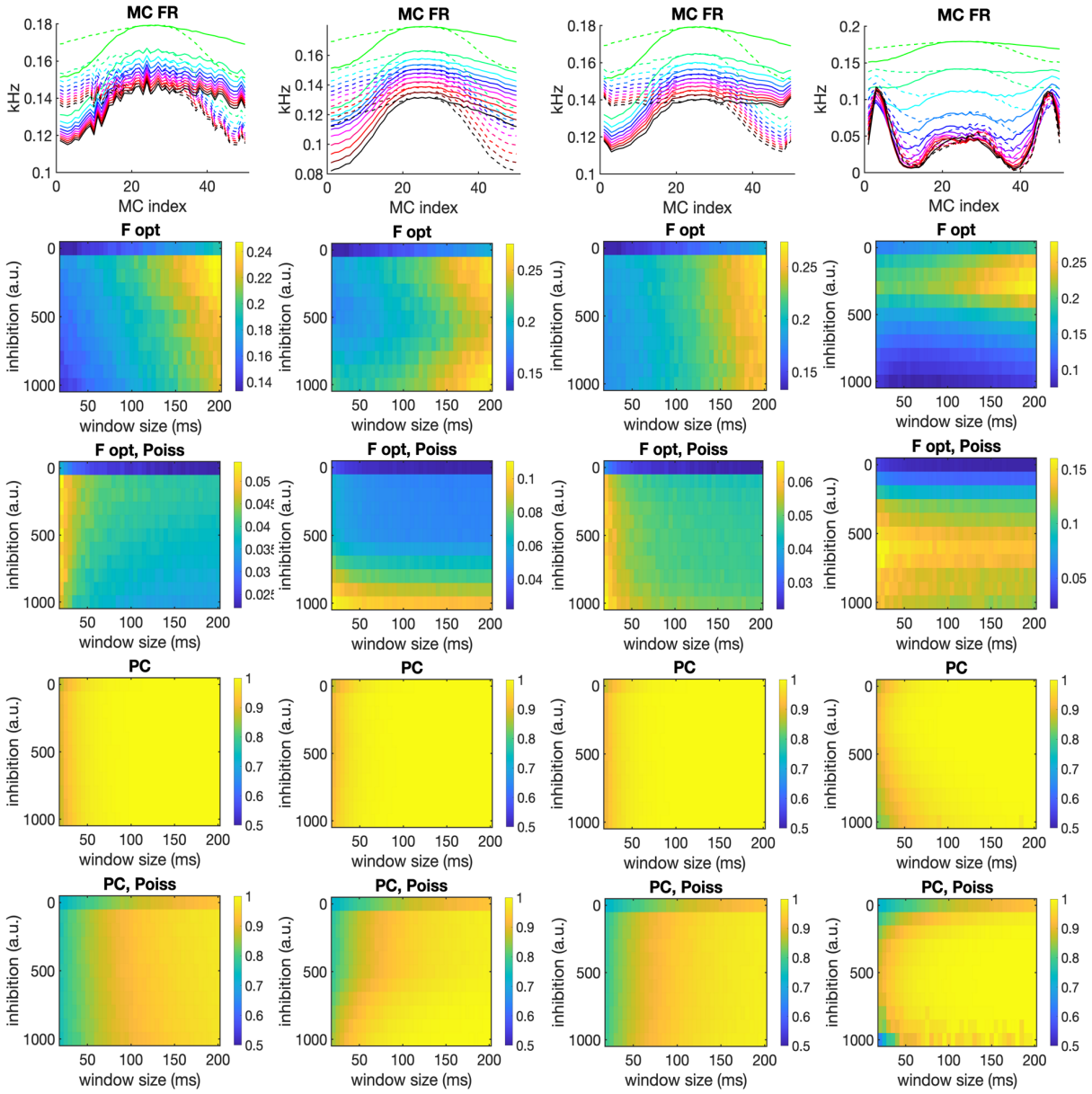


Figure B.22. Same as for B.21 but for skewed Gaussians.

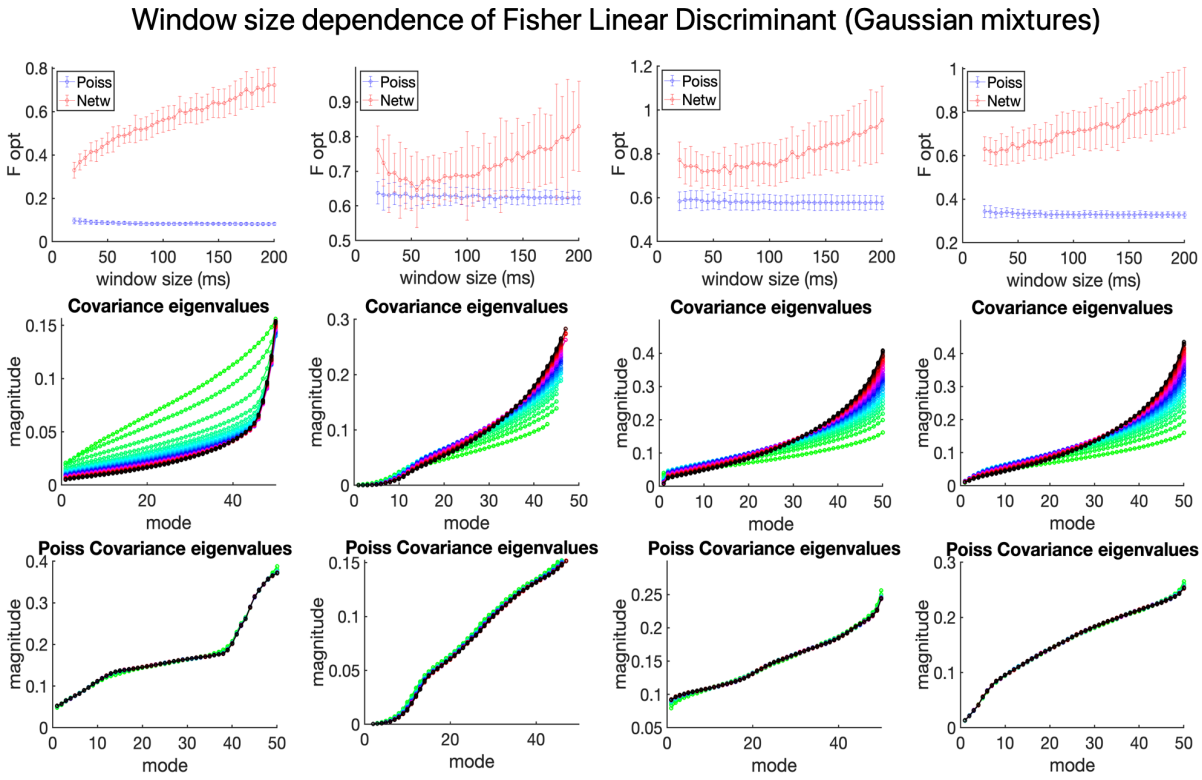


Figure B.23. Columns from left to right indicate single, global, stimulus-based, and Mexican hat connectivity, respectively. Top row indicates $\mathcal{F}_{opt}$ at various measurement bins for network dat (red) and Poisson (blue). Middle row plots eigenvalues of the covariance matrix by ascending magnitude. Bottom row plots eigenvalues of the covariance matrix for the corresponding Poisson network. Green to black indicate increasing inhibition.

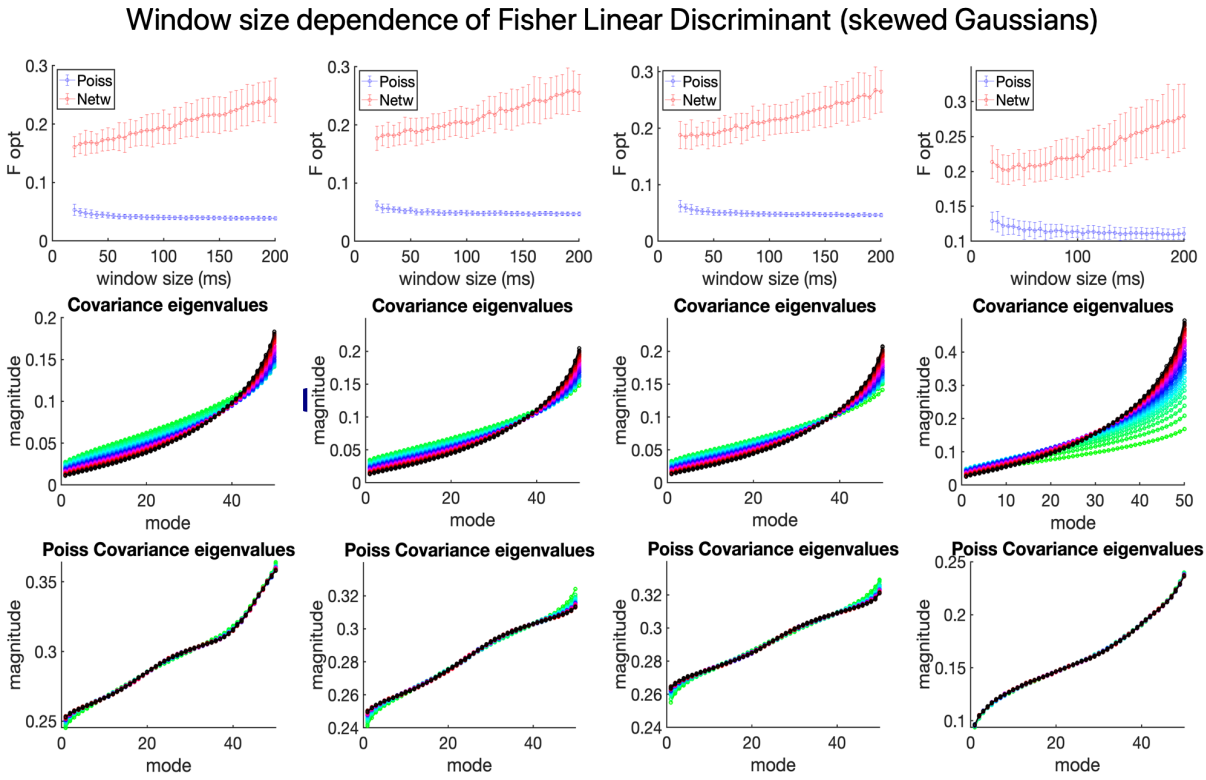


Figure B.24. Same as for B.23 but for skewed Gaussians.

## B.7. Noisy neuron model

The earliest iteration of this project used a stochastic firing neuron, whose implementation is described in [**79**]. These efforts comprised an early work towards understanding networks' roles in shaping noise correlations and thus aiding discrimination.

Consider a simple stochastic neuron model whose voltage-like variable dictates its probability of firing,

$$P\left(u_i\right) = \min\left(\max\left(0, u_i\right), 1\right) \tag{B.105}$$

Its value is the combination of a stimulus, input from recurrent synapses, and a refractory variable,

$$u_i = \eta_i + h_i + \sum_j h_{j,int}, \tag{B.106}$$

where $\eta_i$ is a refractory variable updated as

$$\tau\dot{\eta}_i = -\eta_i \tag{B.107}$$

$$\text{at spike, } \eta_i \to -\eta_0. \tag{B.108}$$

and $h_i$ and $\sum_j h_{j,int}$ comprise external stimuli and internal recurrent inputs,

$$\tau_s\dot{h}_{j,int} = -h_{j,int} \tag{B.109}$$

$$h_{j,int} = h_{j,int} \pm w_{e,i}. \tag{B.110}$$

Dots indicate time derivatives, and subscripts $e, i$ represent excitatory and inhibitory connections, respectively.

We plot firing rates of a 10-MC single-connected network driven by each of two similar odors and with an arbitrary inhibitory strength. Across 100 simulations, cells show reasonably consistent firing patterns, with higher-excited MCs firing more frequently (Figure B.25). In a Poisson network with the same average rates, cells demonstrate less consistency between simulations (Figure B.26).

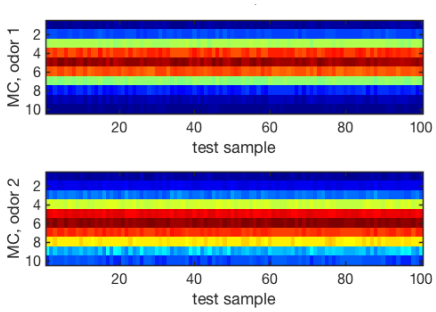


Figure B.25. When spikes are network-driven, mitral activity is consistent across test samples.

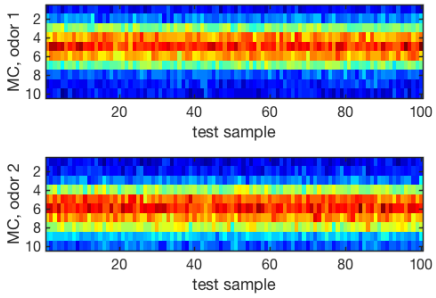


Figure B.26. For Poisson-driven spiking, mitral cells fire less consistently across test samples.

Scatterplots of network-driven firing rates and Poisson rates similarly show that the latter has more variability (Figures B.27 and B.28) and worse discriminability for any inhibitory coupling (Figures B.29 and B.30).

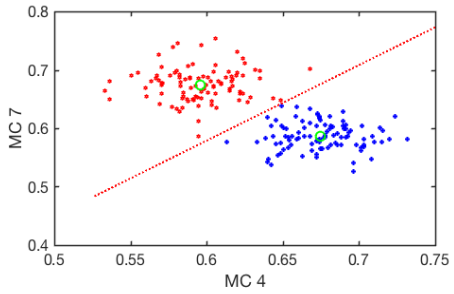


Figure B.27. Activity clouds in a Poisson process have more spread and overlap.

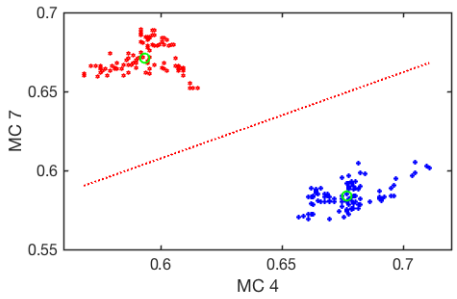


Figure B.28. Activity clouds in a network-spiking process are more compact.

We next assess the eigenmodes of the covariance matrix in each scenario. For Poisson-driven spiking, the magnitudes of eigenvalues scales linearly with the measurement bin, In contrast, for network-driven spiking, eigenvalues of the covariance reduce with window size, even after accounting for the expected scaling (Figures B.31 and B.32). This suggests that recurrent inhibition reduces variance more strongly than what is expected from an

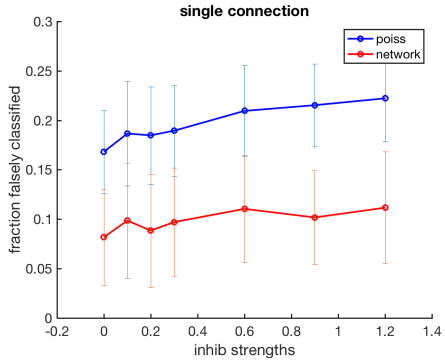


Figure B.29. In a single-connection network, increasing the amount of inhibition increases the rate of false classifications. The effect is slightly larger for Poisson-driven spiking.

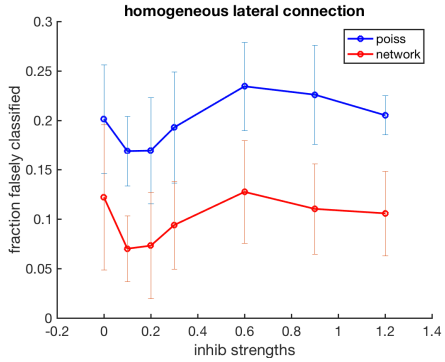


Figure B.30. In a network with all-to-all reciprocal connections between MCs and GCs, there is an optimal amount of inhibition for which linear discrimination is best. Increasing inhibition past that point leads to worse performance.

increase in measurement time, and is a result that is consistent across all our neuron models.

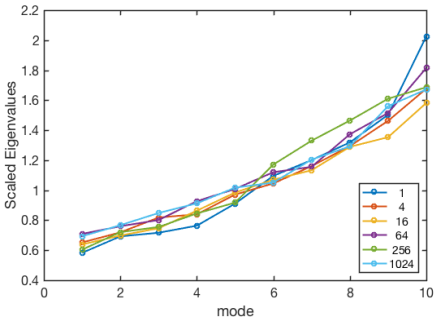


Figure B.31. Eigenvalues (scaled with measurement bin) do not substantially change with measurement bin.

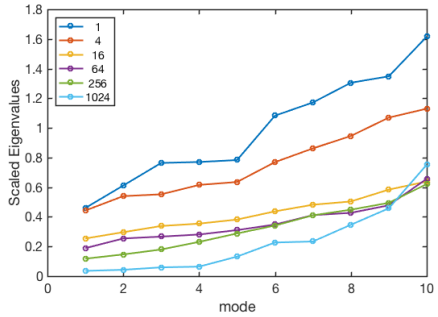


Figure B.32. Scaled eigenvalues of the covariance matrix decrease as window size increases.