Neuromodulatory Effects on Early Visual Signal Processing

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Dominic Gonschorek

aus Berlin

Tübingen

2023

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	18.12.2023
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter/-in:	Prof. Dr. Thomas Euler
2. Berichterstatter/-in:	Prof. Dr. Jan Benda

Erklärung

Ich erkläre, dass ich die zur Promotion eingereichte Arbeit mit dem Titel

„Neuromodulatory effects on early visual signal processing"

selbständig verfasst, nur die angegebenen Quellen und Hilfsmittel benutzt und wörtlich oder inhaltlich übernommene Stellen als solche gekennzeichnet habe. Ich versichere an Eides statt, dass diese Angaben wahr sind und dass ich nichts verschwiegen habe. Mir ist bekannt, dass die falsche Abgabe einer Versicherung an Eides statt mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft wird.

Declaration

I hereby declare that I have produced the work entitled

"Neuromodulatory effects on early visual signal processing",

submitted for the award of a doctorate, on my own (without external help), have used only the sources and aids indicated and have marked passages included from other works, whether verbatim or in content, as such. I swear upon oath that these statements are true and that I have not concealed anything. I am aware that making a false declaration under oath is punishable by a term of imprisonment of up to three years or by a fine.

Tübingen, den ....
Datum/Date Unterschrift/Signature

Contents

Abstract

Understanding how the brain processes information and generates simple to complex behavior constitutes one of the core objectives in systems neuroscience. However, when studying different neural circuits, their dynamics and interactions researchers often assume fixed connectivity, overlooking a crucial factor - the effect of neuromodulators. Neuromodulators can modulate circuit activity depending on several aspects, such as different brain states or sensory contexts. Therefore, considering the modulatory effects of neuromodulators on the functionality of neural circuits is an indispensable step towards a more complete picture of the brain's ability to process information. Generally, this issue affects all neural systems; hence this thesis tries to address this with an experimental and computational approach to resolve neuromodulatory effects on cell type-level in a well-define system, the mouse retina.

In the first study, we established and applied a machine-learning-based classification algorithm to identify individual functional retinal ganglion cell types, which enabled detailed cell type-resolved analyses. We applied the classifier to newly acquired data of light-evoked retinal ganglion cell responses and successfully identified their functional types. Here, the cell type-resolved analysis revealed that a particular principle of efficient coding applies to all types in a similar way.

In a second study, we focused on the issue of inter-experimental variability that can occur during the process of pooling datasets. As a result, further downstream analyses may be complicated by the subtle variations between the individual datasets. To tackle this, we proposed a theoretical framework based on an adversarial autoencoder with the objective to remove inter-experimental variability from the pooled dataset, while preserving the underlying biological signal of interest.

In the last study of this thesis, we investigated the functional effects of the neuromodulator nitric oxide on the retinal output signal. To this end, we used our previously developed retinal ganglion cell type classifier to unravel type-specific effects and established a paired recording protocol to account for type-specific time-dependent effects. We found that certain retinal ganglion cell types showed adaptational type-specific changes and that nitric oxide had a distinct modulation of a particular group of retinal ganglion cells.

In summary, I first present several experimental and computational methods that allow to study functional neuromodulatory effects on the retinal output signal in a cell type-resolved manner and, second, use these tools to demonstrate their feasibility to study the neuromodulator nitric oxide.

Zu verstehen wie das Gehirn Informationen verarbeitet und einfaches bis komplexes Verhalten erzeugt, ist eines der zentralen Ziele der systemischen Neurowissenschaften. Bei der Untersuchung verschiedener neuronaler Schaltkreise, ihrer Dynamik und ihrer Interaktionen gehen Forscher jedoch häufig von einer festen Konnektivität aus und übersehen dabei einen entscheidenden Faktor - die Wirkung von Neuromodulatoren. Neuromodulatoren können die Aktivität von Schaltkreisen in Abhängigkeit von verschiedenen Aspekten modulieren, z.B. von unterschiedlichen Gehirnzuständen oder sensorischen Kontexten. Die Berücksichtigung der modulierenden Wirkung von Neuromodulatoren auf die Funktionalität neuronaler Schaltkreise ist daher ein unerlässlicher Schritt auf dem Weg zu einem vollständigeren Bild der Informationsverarbeitungsfähigkeit des Gehirns. Generell betrifft dieses Problem alle neuronalen Systeme; daher wird in dieser Arbeit versucht, dieses Problem mit einem experimentellen und computergestützten Ansatz zu lösen, um neuromodulatorische Effekte auf Zelltyp-Ebene in einem gut definierten System, der Mausretina, zu klären.

In der ersten Studie haben wir einen auf maschinellem Lernen basierenden Klassifizierungsalgorithmus entwickelt und angewendet, um einzelne funktionelle retinale Ganglienzelltypen zu identifizieren, was detaillierte zelltypaufgelöste Analysen ermöglichte. Wir waren in der Lage, den Klassifikator auf neu gewonnene Daten von lichtevozierten retinalen Ganglienzellantworten anzuwenden und ihre funktionellen Typen erfolgreich zu identifizieren. Dabei zeigte die zelltypaufgelöste Analyse, dass ein bestimmtes Prinzip der effizienten Kodierung für alle Typen in ähnlicher Weise gilt.

In einer zweiten Studie konzentrierten wir uns auf das Problem der interexperimentellen Variabilität, die bei der Zusammenführung von Datensätzen auftreten kann. Daher können weitere nachgelagerte Analysen aufgrund der subtilen Unterschiede zwischen den einzelnen Datensätzen erschwert sein. Um dieses Problem zu lösen, haben wir einen theoretischen Ansatz vorgeschlagen, der auf einem adversen Autoencoder basiert und darauf abzielt, die interexperimentelle Variabilität aus dem gepoolten Datensatz zu entfernen, während das zugrunde liegende biologische Signal erhalten bleibt.

In der letzten Studie dieser Arbeit untersuchten wir die funktionellen Auswirkungen des Neuromodulators Stickstoffmonoxid auf das retinale Ausgangssignal. Zu diesem Zweck verwendeten wir unseren zuvor entwickelten Klassifikator für retinale Ganglienzelltypen, um typspezifische Effekte aufzuschlüsseln, und erstellten ein gepaartes Aufnahmeprotokoll, um typspezifische zeitabhängige Effekte zu berücksichtigen. Wir fanden heraus, dass bestimmte retinale Ganglienzelltypen adaptive typspezifische Veränderungen aufweisen und dass Stickstoffmonoxid eine deutliche Modulation einer spezifischen Gruppe der retinalen Ganglienzellen bewirkt.

Zusammenfassend stelle ich zunächst mehrere experimentelle und computergestützte Methoden vor, die es ermöglichen, funktionelle neuromodulatorische Effekte auf das retinale Ausgangssignal zelltypspezifisch zu untersuchen, und zeige dann anhand dieser Instrumente ihre Durchführbarkeit bei der Untersuchung des Neuromodulators Stickstoffmonoxid.

Synopsis

Neuromodulation: A Wireless Network

The human brain is one of the fundamental elements of our identity. It enables us to perceive the world, process information, have subjective experiences, be self-aware, as well as possess other complex cognitive abilities. For centuries, humankind tried to gain insights into this biological machinery in all its complexity to understand themselves, but also to advance in several other aspects of human life, such as curing diseases.

The fundamental building blocks of the brain are the countless sophisticated neural circuits that process inconceivable amount of information to generate various behaviors. To unravel the functionality of these circuits, one needs to consider their individual components, their connections, and interactions with other circuits. Over decades, scientists across all disciplines of neuroscience sought to study the neural circuits throughout the nervous system ranging from the associated peripheral and sensory systems to higher brain areas. Studying and understanding neural circuits in any brain area is very complex as the information input to each neuron and its output is difficult to define due to the high degree of interconnectivity and feedback mechanisms. Here, a keystone direction of research constitutes 'connectomics', which aims to elucidate the complexity of neural circuits by mapping entire connections between every neuron and every synapse of the nervous system (Seung, 2012). Despite this pioneering work, a crucial aspect that is often neglected when studying the functionality of neural circuits, yet adds additional layers of complexity, is the influence of neuromodulators, which can alter and regulate neurophysiological processes (Bargmann, 2012; Marder, 2012). For example, they have implications and regulatory mechanisms of different functions within the brain such as learning and motor control (Volkow and Morales, 2015), and their dysfunction may be the cause of several psychiatric disorders (Yang and Tsai, 2017; Meder et al., 2019). Therefore, deepening our knowledge of neuromodulators across neural systems will help us to address various issues from cellular mechanisms to curing diseases. Since neuromodulators may affect different circuits within a neural system differentially, the resolution down to cell type-level constitutes an important step towards a detailed functional understanding of neuromodulatory implications. In the first part of this thesis, I elucidate why the retina serves as an ideal model system for studying neuromodulation and highlight the development of a retinal ganglion cell (RGC) type classifier to functionally identify the vast diversity of the retinal output. In the second part, I address the issue of inter-experimental variability between recordings of different experiments and datasets and how this may cause false-positive interpretations, especially when studying neuromodulatory effects. Here, I provide two possible solutions: *(i)* a theoretical framework and *(ii)* an experimental approach. In the final part of my thesis, I leverage the developed tools in a study to showcase the neuromodulatory effects of nitric oxide (NO) on various RGC types, while considering inter-experimental variability to disentangle experimentally-induced from drug-induced effects.

The Mouse Retina as a Model System for Vision

The choice of an ideal model system to study a particular research question is crucial for the outcome, interpretation, and translation of the results to other systems. In the context of studying the functional effects and implications of a neuromodulator on a well-defined system, the mouse retina constitutes a suitable system for several reasons. As part of the central nervous system, the retina has the advantage that we can experimentally define the neural input to the system by displaying light stimuli and directly measuring its neural output signal to the brain with several physiological techniques. The mammalian retina barely receives any external input or feedback from the brain (with minor exceptions (Gastinger et al., 2006a)), other than within the rest of the central nervous system, where feedback mechanisms and undefined additional inputs from other brain regions contribute to the overall signal. This characteristic enables the experimenter to manipulate and isolate specific parameters using neuromodulators to infer their effects on the system. Moreover, the retina releases neuromodulators endogenously; thus providing an inherent system to study their effects on information processing. In this section, I will first explain the structure of the mouse retina, how the signal of the visual environment is processed along the retinal signaling pathways, and lastly, elaborate the vast repertoire of neuromodulators in the retina.

Early Visual Signal Processing

For an organism to perceive its surrounding environment, it must be able to process various kinds of information. To this end, organisms developed several sensory system to sense their environment in order to navigate, hunt, mate, and simply survive. One of the primary sensory systems for many animals, including humans, is the visual system. To perceive the visual environment, photons are captured by a thin sheet of neuronal tissue at the back of the eye, termed the 'retina'.

One of the major challenges for the retina is to capture the complexity of the visual environment in all its ecologically relevant facets as a 'neural image'. Such an image of the natural environment encodes various visual features such as contours, contrast, light intensity, luminance, color, and many more. Besides that, the visual environment is not static, but dynamic, so the retina needs to additionally capture and process information about the temporal domains, such as motion, changes in luminance and other changing image statistics (Dong and Atick, 1995; Olshausen and Field, 1996; Van Hateren and Schaaf, 1998). Impressively, despite having a rather simple layered structure and comprising only five main cell classes, its actual complexity is based on the vast diversity of neuron types within each of these classes (Wässle and Boycott, 1991; Masland, 2001; Wässle, 2004; Masland, 2012a). Using this diversity, the retina forms sophisticated and specialized circuits to parallelize the information processing as well as to operate at different light conditions throughout the day (Dacey, 2004). Notably, specific parts of circuits are used for various purposes and processing steps to further broaden the operational capabilities of the retina (Smith et al., 1986). But how does the information flow through the retinal network and what are the different processing stages as soon as a stream of photons hits the first sensory neurons within the retina?

In brief, a stream of photons is captured and converted into an electrical signal by the first sensory neurons, the photoreceptors (PRs), which traverse the signal via bipolar cells (BCs) to the output neurons, the retinal ganglion cells (RGCs) (Masland, 2012a; Euler et al., 2014). Then, RGCs sent the signal to various higher brain areas as spike trains (Dhande and Huberman, 2014; Seabrook et al., 2017). The signal, which is sent along this vertical,

excitatory-driven pathway, is additionally shaped by two lateral, inhibitory neuron classes located in the outer and inner retina, namely horizontal cells (HCs) and amacrine cells (ACs), respectively.

In depth, the first stage of the processing of visual information in the retina constitutes the event of photon capturing by PRs, which are located in the outer part of the retina, specifically in the outer nuclear layer (ONL). Within the mouse retina, there are three types of photoreceptors specialized for different tasks, namely rods, cones and intrinsically-photosensitive RGCs (ipRGCs), also called melanopsin-containing RGCs (mRGCs). Importantly, the latter type forms the basis for non-image forming vision and is not located within the ONL, but in the ganglion cell layer (GCL) and mainly contributes to the synchronization of the circadian rhythms (Freedman et al., 1999; Berson et al., 2001; Hattar, 2003), regulation of the pupil size (Chen et al., 2011) and recently was shown to be a ground luminance detector (Berry et al., 2022). In contrast to ipRGCs, rods and cones constitute the basis for image-forming vision. Both types convert the stream of photons into an electrical signal via an intrinsic phototransduction cascade (Chen, 2005; Yau and Hardie, 2009) and incorporate different tasks to operate at distinct light conditions with overlapping signalling pathways. While rods function under scotopic (dim) light conditions due to their high sensitivity, cones, being less sensitive, operate with increasing luminance towards photopic (bright) conditions. This is mainly due to their intrinsic properties as rod responses saturate at high luminance levels, whereas then cones take over as they do not saturate, however, their operational response range is still limited by their photoisomerization rate (Baylor, 1987). In mice, cones express distinct wavelength sensitivities by expressing two types of opsins (S- and M-opsins), whereby S-opsin exhibits a peak sensitivity to short ('S') wavelength at 360 nm and M-opsins to medium ('M') wavelength at 508 nm (Nikonov et al., 2006). In addition, the mouse retina features a specific adaptation by expressing 'true' S-cones (exclusively expressing S-opsin) in the upper visual field and M-opsins co-expressing S-opsins towards the lower visual field (Applebury et al., 2000; Nadal-Nicolás et al., 2020). These features of cones form the first stage of color detection.

After the conversion of photons into an electrical signal, the signal is forwarded to BCs. Here, the signal is shaped by a network of electrically coupled HCs, which provide a variety of inhibitory feedback mechanisms by forming direct synaptic connections to rods and cones (Masland, 2001; Kamermans and Spekreijse, 1999; Thoreson and Mangel, 2012). These feedback mechanisms act locally and globally (Behrens et al., 2022) and serve several functions such as contrast enhancement, light adaptation, as well as the formation of a center-surround receptive field structure (Diamond, 2017; Ströh et al., 2018; Drinnenberg et al., 2018). Together, this network of synaptic connections of PRs, HCs and BCs form the outer plexiform layer (OPL), which traverses the signal from the ONL (rods and cones) to the inner nuclear layer (INL; HCs, BCs and ACs). From here, the signal divergences and is parallelized into different feature channels since rods and cones connect to different BC types (Euler et al., 2014; Behrens et al., 2016). In mice, the signals transmitted by cones are relayed to 13 different BC types, the cone BCs (CBCs), whereas rod signals are only relayed to one BC type, the rod BC (RBC) (Wässle et al., 2009; Shekhar et al., 2016). The signal divergence and parallelization is based on the different attributes and features of the BC types, resulting in diverse response patterns (DeVries, 2000; Wässle et al., 2009; Baden et al., 2013). In general, BC types can be distinguished based on several characteristics: protein expression, axonal stratification depth in the sublamina of the inner plexiform layer (IPL), and their dendritic and axonal morphology (Euler et al., 2014). Functionally, BC types can be split into On- and Off-BC types based on their response polarity, i.e., depolarization in response to light increments or decrements, respectively (DeVries, 2000). This response polarity depends on the expression of glutamate

receptors at their dendrites: On-BCs express the metabotropic receptor mGluR6, resulting in a sign-inverting signal transmission, whereas Off-BCs express sign-conserving ionotropic receptors (kainate and/or AMPA). Furthermore, functional BC types differ in their temporal response kinetics to light illumination (Baden et al., 2013), and can be divided in sustained or transient response types. Additionally, certain BC types carry chromatic information, whereas others are achromatic (Breuninger et al., 2011; Euler et al., 2014). Interestingly, within the mammalian retina, BC types have a distinct stratification profile of their axon terminals in the sublamina of the IPL with Off-types stratifying in the upper and On-types towards the bottom layers (Wu et al., 2000; Ghosh et al., 2004).

Here, neuron classes of the INL, mainly BCs and ACs, form a dense network of synaptic connections with dendrites of RGCs, which are located at the innermost layer of the retina, the GCL. Within the IPL, the excitatory glutamatergic input from BCs is further transferred to various RGCs types. With more than 40 types in mice (Baden et al., 2016; Goetz et al., 2022), the functional diversity of RGCs enables the retina to further process the input signal, perceived by PRs, in parallel and to extract a broad variety of relevant features of the visual world. The extracted information is then relayed via their axonal projections out of the retina to different visual brain areas (Seabrook et al., 2017). This functional diversity of RGC types is based on at least three factors: *(i)* different intrinsic properties such as receptor expression (Wienbar and Schwartz, 2022), *(ii)* distinct and specialized connections with BC types (Nirenberg and Meister, 1997; Asari and Meister, 2012), and *(iii)* selective connections with ACs types (Gollisch and Meister, 2010; Masland, 2012b). Especially the latter ones additionally shape and modulate the excitatory input from BCs to RGCs (Diamond, 2017). ACs represent the most diverse and complex class of neurons within the retina. To date, more than 60 molecularly types were identified (in mice), however, little is known about many of them (Yan et al., 2020). There are only few well-studied types, such as the AII (Demb and Singer, 2012; Marc et al., 2014) and starburst AC (Euler et al., 2002; Masland, 2005; Taylor and Smith, 2012), for which their function and mechanism(s) within the retinal circuitry is identified. In general, ACs have a broad repertoire of mechanisms to increase the computational capacities of the retina. They provide feedforward inhibition to RGCs, feedback inhibition to BCs as well as lateral inhibition to other ACs, which are mainly mediated by their primary inhibitory neurotransmitter, GABA (wide-field ACs) or glycine (small-field ACs) (Diamond, 2017). Importantly, ACs can be considered to be 'bilingual': aside of their fast-acting inhibitory neurotransmitter, they express either another fast-acting neurotransmitter such as glutamate (e.g., VGluT3 ACs (Haverkamp and Wässle, 2004)) or one of the many slow-acting neuromodulators, e.g., nitric oxide (Kim et al., 1999; Jacoby et al., 2018) or dopamine (Dacey, 1990). This breadth and diversity of possibilities expands the operational capability and computations of the retina to process nearly all conditions of the visual environment that are relevant to the organism. In the next section, I aim to shine light on the functionality and diversity of endogenously occurring neuromodulators in the mouse retina.

A Diverse Cocktail of Neuromodulators

The processing of information within a neural system, such as the mouse retina, is primarily driven by excitatory and inhibitory neurotransmitters, e.g., glutamate and glycine/GABA, respectively. Here, the excitatory signals, transmitted by glutamate, are sent from PRs to RGCs via BCs and is additionally shaped by the inhibitory input from HCs and ACs, releasing glycine and GABA. However, to adapt to different sensory and contextual states, the retina

uses another group of molecules that affect the signal processing on a rather slow-acting timescale, namely neuromodulators (Diamond, 2017). Unlike neurotransmitters, neuromodulators modulate the transmitted signal and activity of an existing neural circuit (Bargmann, 2012). Therefore, they broaden and extend the functionality of 'hard-wired' neural circuits depending on the sensory context by affecting multiple neurons and synapses within such a circuit.

Neuromodulators feature a range of distinct mechanisms to modulate the activity within a neural circuit. For example, they can directly or indirectly activate and bind to receptors or ion channels, triggering intracellular signaling cascades that influence neuronal excitability and functionality of synaptic transmission (Harvey and Heinbockel, 2018). Also, neuromodulators can regulate the strength of a signal and its synaptic transmission by amplifying or dampening the neurotransmitter release, their sensitivity or reuptake into the synapse (Nishizaki et al., 2002). Another key mechanism constitutes retrograde signaling in which the activity of the presynaptic neuron is modulated by a neuromodulator released from the postsynapse (Alger, 2002). Notably, they can also regulate feedback mechanisms of their own release or even of others to tune and configurate various circuits (Jacoby et al., 2018). By this, neuromodulators provide a broad spectrum of mechanisms on how to influence the activity of neural circuits.

Overall, the retina contains and releases a vast and diverse repertoire of neuromodulators, comprising at least 20 different types (Diamond, 2017). This diversity ranges from monoamines (e.g., histamine (Gastinger et al., 2006a), dopamine (Witkovsky and Dearry, 1991)), endocannabinoids (Yazulla, 2008), gasotransmitters such as nitric oxide (NO) (Goldstein et al., 1996) and carbon monoxide (CO) (Cao et al., 2000), and neuropeptides (e.g., neuropeptide Y (Santos-Carvalho et al., 2014; Santos-Carvalho et al., 2015)). As mentioned earlier, the majority of these neuromodulators is released by subtypes of ACs in additional to their primary neurotransmitter, whereas small fractions of them is released outside the retina by centrifugal fibers, e.g., serotonin and histamine (Gastinger et al., 2006b). In particular, serotonin is released from neuronal projections originating in the dorsal raphe and histamine from projections in the hypothalamus. Notably, both have been associated with the arousal system, thereby modulating the retinal code depending on the arousal state of the animal (Schröder et al., 2020b; Warwick et al., 2022). Nevertheless, only a few neuromodulators are partly understood, yet a comprehensive and integrative picture is slowly emerging. For example, dopamine constitutes one of the better-studied neuromodulators, still the overall picture is puzzling, despite knowing much about its release site, circuit mechanisms, and involved cell types (Roy and Field, 2019; Warwick et al., 2023). Dopamine is released from a specialized ACs within the inner retina, the dopaminergic AC (DAC), which receives input from various cell types such as CBCs, ipRGCs, AII ACs and other ACs (Zhang et al., 2008; Prigge et al., 2016; Qiao et al., 2016; Zhao et al., 2017). Interestingly, DACs co-release GABA, and ATP, in addition to dopamine (Ho et al., 2015; Hirasawa et al., 2015), whereby the release of these three molecules is triggered by different mechanisms at different release sites along the dendrites (Pérez-Fernández et al., 2019). Dopamine has been implicated and shown to exhibit a broad variety of functionality in the retina such as light adaptation (Witkovsky, 2004), regulation of the circadian rhythm (Nir et al., 2000), and modulation of retinal processing (Roy and Field, 2019). Despite knowing the release sites, receptors and cellular mechanisms for various neuromodulators, a comprehensive, systematic, and cell type-resolved overview of their function on visual signal processing is yet missing.

As a consequence, I sought to tackle this problem by developing an experimental pipeline, in which we established a RGC type classifier to break down the retinal code in its individual components and addressed additional issues when studying neuromodulatory effects. In the next chapter, I provide an overview of approaches to classify RGC types and introduce our RGC type classifier.

Functional Cell Type Classification

In the first part of my thesis, I address the need and development of an RGC type classifier that is easy to use, yet precise in its prediction to map RGCs to their the functional types. To understand how the functionality of the retinal code, represented by its output neurons, is modulated by neuromodulators, the disentanglement of the retinal code down to cell type-level is necessary, otherwise a detailed, yet broad picture of their implications for vision cannot emerge. In this section, I explain how RGC types were usually classified, how a large-scale clustering of RGCs set new standards, followed by previous approaches and how our RGC type classifier makes it possible to identify functional RGC types using only two short and simple stimuli.

Approaches to Classify RGC Types

To unravel the retinal code, it is essential to dissect it into its individual components due to the parallelization of the signal into different feature channels. Therefore, the identification and classification of the different functional RGC types is a necessary consequence. Such a classification of RGC types serves multiple purposes: *(i)* it allows the identification of individual types and their contribution to specific features and aspects of the visual information processing, *(ii)* it helps to understand how the visual signal is encoded and processed by the diversity of RGC types, *(iii)* it allows to disentangle sub-circuits, and *(iv)* it helps understanding retinal diseases that may affect specific and individual RGC types.

RGC types were distinguished based on three main criteria: *(i)* their morphological characteristics, *(ii)* their genetic profile, and *(iii)* their functional properties (Sanes and Masland, 2015; Vlasits et al., 2019). The morphological classification makes use of different microscopy techniques, such as light microscopy, confocal microscopy, and electron microscopy (EM), to reconstruct and characterize the dendritic and axonal structures of individual RGC types. However, this does not provide any information about their physiology and is subject to sampling bias. Using these tools, earlier studies identified approx. 20 RGC types (Völgyi et al., 2009), whereas a more recent study was able to identify >40 anatomical types (Bae et al., 2018). On the other hand, genetic and molecular approaches aim to classify RGC types based on their expression of specific genetic markers. Here, researchers make use of techniques such as immunohistochemistry to identify these markers, transgenic mouse lines to visualize and study distinct RGC types expressing certain genes, and transcriptomics to assess and analyze their gene expression patterns. The drawback of these techniques is that they are usually slow and only target one type or a particular subgroup of types, making it difficult to grasp the complete diversity of RGC types. With advances in technology, novel high-throughput tools, such as single-cell transcriptomics, found >40 distinct RGC types (Tran et al., 2019).

Since the overall functional RGC classification and typing may be close to complete, in this thesis, I focus on the mapping of existing functional types to newly acquired data to simplify their identification in the form of a supervised classifier. Therefore, I emphasize the functional classification of RGCs in the remaining part of this section.

Functionally, RGC types were mainly classified based on their light-evoked response properties to identify how and which features of the visual world are encoded by distinct RGC types (Lettvin et al., 1959; Werblin and Dowling, 1969). With the discovery of center-surround receptive fields in certain RGCs by Barlow and Kuffler, they revealed a fundamental principle and set the foundation for understanding the physiological basis of different RGC types (Barlow et al., 1957). This principle was extended by the findings from Hubel and

Wiesel when they identified two types of RGCs (X and Y cells) that exhibited distinct response properties of their spatial receptive fields and preferences to different stimulus features (Hubel and Wiesel, 1959; Hubel and Wiesel, 1962). Based on these results, researchers started to further distinguish types based on their response to light increments and decrements, which resulted in the classification of ON, OFF or ON-OFF types (Levick, 1967). Additionally, the temporal response kinetics constituted a crucial feature, meaning if they exhibit a sustained or transient response to either light increment or decrement (Cleland et al., 1971).

The discovery that RGCs project to various brain regions revealed parallel processing pathways, which suggested an even broader diversity of RGCs (Roska and Werblin, 2001). Taking this into account, the spectrum of features to include for the classification broadened and resulted in a functional diversity of RGCs comprising >40 types to date (Baden et al., 2016; Goetz et al., 2022). This revealed the identification of distinct cell types, which contribute to specific visual processes such as direction-selectivity (Weng et al., 2005), local-edge detectors (Van Wyk et al., 2006), and color-opponent cells (Szatko et al., 2020).

Nowadays, two main techniques are used for functional recordings: (i) electrophysiological tools, such as patch clamp or extracellular recordings (multi-electrode recordings (MEA)) (Grumet et al., 2000), and (ii) optical methods, such as two-photon calcium imaging (Euler et al., 2019). Among those techniques, the high-throughput recordings (MEA and two-photon calcium population imaging) of RGC populations allowed researchers to gain a comprehensive overview of the functional diversity of RGC types. Both techniques have their advantages and disadvantages. For example, MEA recordings allow measuring populations of RGCs with a high temporal resolution, yet it does not provide any information about the morphology and is prone to sampling bias as cells with no to few spikes are not recognized (Trapani et al., 2022). Functional calcium imaging, on the other hand, solves this issue by exciting the inserted or expressed fluorescent indicators in each cell, recording the changes of calcium activity and therefore making each cell in the recording field visible (Euler et al., 2019). A downside of this is that calcium only serves as a proxy for spiking activity, and the relationship between spikes and calcium is nonlinear, which may result in potential underestimations of specific feature representations (Kerr et al., 2000; Vogelstein et al., 2010; Baden et al., 2016). Additionally, calcium changes and the indicator kinetics are much slower than actual spikes; thus certain rapid changes in spiking activity cannot be resolved. Despite their disadvantages, the overall results leveraged the classification of functional RGCs in a direction where the picture starts to get complete. Further technical advances, such as voltage imaging, may enable us to resolve the functional diversity of RGCs in an even more detailed way to complement missing pieces and information (Knöpfel and Song, 2019; Liu et al., 2022).

In this context, over the last decades, many studies aimed to combine and align the different classification approaches to provide a coherent and unified picture of the diversity of RGC types in the mouse retina (Seung and Sümbül, 2014; Kölsch et al., 2021). Recently, Goetz et al. (2022) addressed this problem by aligning a large dataset of functional RGC responses with their morphological and genetic markers. Using these three features in a joint approach, they were able to classify 42 RGC types. With advancing techniques, we will be able to gain more insights into the diversity of RGC types, including even more features to further understand their implications for vision.

In the next subsection, I summarize the results of our supervised RGC type classifier and how a study from 2016 set the foundation for the current classification approach to map existing types to new data in order to identify RGC types.

RGC Type Classifier: A Supervised Learning Approach

In 2016, Baden et al. sought to unravel the functional diversity of RGC types in the mouse retina and set a new standard for such cell type classifications. To this end, they used a high-throughput approach by performing two-photon calcium imaging of RGC populations to record their light-evoked activity. To capture the vast diversity of RGC types, they used a battery of artificial stimuli, e.g., moving bar and full-field chirp, that covered a broad range of the stimulus feature space. This way, they increased the likelihood of evoking individual response preferences of each type. Using an unsupervised clustering approach (Gaussian mixture model) of the dimensionality reduced response features, and additional features e.g., soma size and direction-selectivity index, they identified 32 functional RGC and 14 displaced AC types within the GCL.

Since then, this set a gold standard for functional RGC types and several studies used them as reference. For example, a publication from 2019 recorded RGC and dLGN responses to map individual types between these areas (Rosón et al., 2019). To do so, they used a correlation-based metric, i.e., 'match index', with which they identified certain RGC types. Furthermore, these reference types were also used to identify RGC types recorded with other physiological tools such as MEA recordings (Shabani et al., 2021; Goldin et al., 2022). To match the reference types, they convoluted the spike trains with the corresponding fluorescent indicator kernel to transform the light-evoked spiking patterns into pseudo-calcium responses. Then, using a correlation-based approach, they identified and matched most RGC types.

To date, the 'match index' is the only approach that maps functional RGC types using two-photon calcium imaging based on their response profiles (Rosón et al., 2019). Nevertheless, this approach had the disadvantage of being biased towards the simple stimulus, i.e., the moving bar in this case, as both used stimuli (full-field chirp and moving bar) were considered to be equally important for the type prediction. Due to the complexity of the chirp, exhibiting a much broader variety of features, the 'match index' was unreliable across types. Moreover, it only included the response quality as an additional parameter, but no other type-specific features such as soma size or direction selectivity, which led to less precise predictions. Since various types show similar responses to chirp and/or moving bar, yet are mainly distinguishable using those features, adding those features would add additional value to the prediction (Baden et al., 2016). As a consequence, there was the need for a reliable and feature-based approach to predict RGC types.

To overcome this, we used a supervised machine learning approach, namely a Random Forest classifier (Ho, 1995), and the previously mentioned dataset comprising the light-evoked responses of the 32 RGC and 14 dAC types (Baden et al., 2016). In particular, we used $\sim$8000 RGC responses to the chirp and moving bar from Baden et al. (2016), projected them into feature space using the extracted visual feature matrices (20 features for the chirp and 8 features for the moving bar). For this, we calculated the dot product between the extracted features and the normalized responses. Additionally, two parameters were used to distinguish alpha and non-alpha cells, as well as direction-selective and non-direction-selective cells by adding the soma size and the p-value of the permutation test for direction selectivity, respectively. Then, the classifier was trained on a subset and validated on a held-out test dataset. After hyperparameter search and cross-validation, it reached a prediction accuracy of $\sim$76% across the 46 GCL types. As output, the classifier returns two parameters: *(i)* an array of 46 confidence scores, i.e., the probability of each type predicted by the multiple decision trees, and *(ii)* the actual predicted RGC type, namely the one with the highest confidence score among all types. This output constitutes an advantage as it provides a probability distribution and not just a single value. Using this, we can adjust and explore the

trade-off between prediction accuracy and fraction of unclassified cells to improve the overall prediction performance of the classifier (Fig. 1b,c). In our case, this yielded an improvement of the prediction performance of the classifier to ∼85% (Fig. 1a). Thus, setting a threshold of 0.3 fulfilled an appropriate balance between these two criteria.

While the original RGC clustering used 40 features in total, including information about receptive fields and color, our RGC classifier only uses 30 features originating from only two stimuli and the other two additional parameters. On one side, this may explain why we were not able to improve the prediction performance of the classifier even further, which may be one key limitation of the classifier. However, on the other side, we aimed to develop a classifier that needs minimal input as usually an experimenter wants to display and record responses to additional stimuli in order to address other scientific questions, and thus only use the two stimuli to identify the types. Again, this constituted a trade-off between achieving a marginally improvement of the prediction and performing efficient experiments. Moreover, prolonged recordings by adding more stimuli for the classification may accelerate the effects of calcium buffering and photobleaching, which may result in a reduction of the signal-to-noise ratio of the responses to other stimuli of interest.

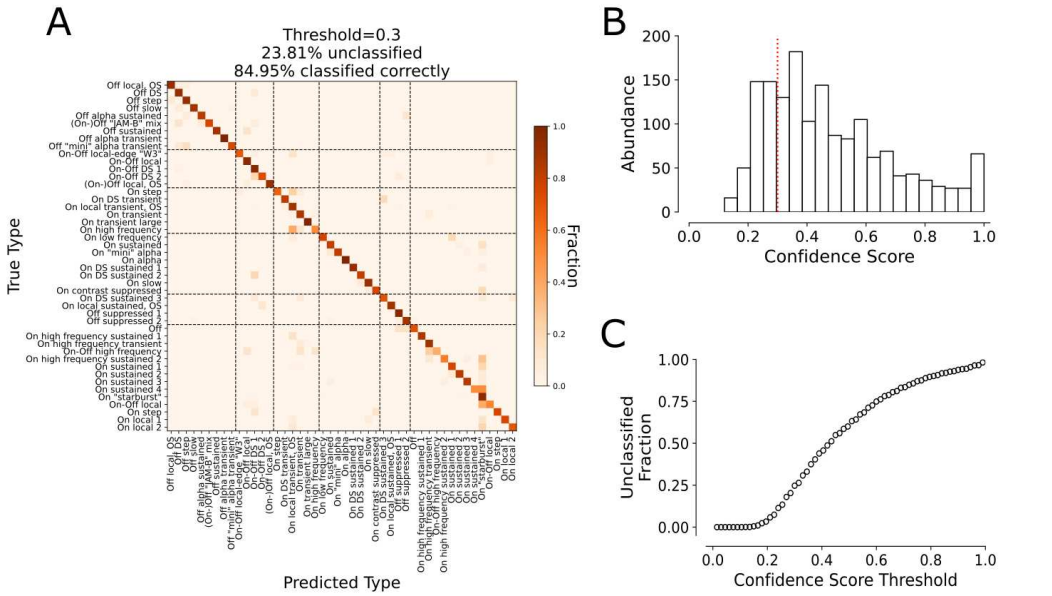

Figure 1: Validation of the RGC type classifier performance. (a) Confusion matrix for the RGC type classifier performance with confidence score threshold of 0.3. Dotted line indicated separation of larger subgroups. Data taken from Baden et al. (2016). **(b)** Confidence score distribution of the highest score per cell of the test data. Red dotted line indicated confidence threshold. **(c)** RGC type classifier confidence score threshold as a function of the fraction of unclassified cells.

In the first study, in which the classifier was used, we developed a modeling approach to predict RGC responses to visual stimuli considering the efficient coding principle and including statistics from natural scenes (Appendix, Study I). We were able to show that these aspects allowed a significant improvement in the predictive performance (Fig. 6, Appendix, Study I). Since we asked whether this principle transfers and affects all RGC types or only particular types, we used the RGC type classifier to identify the individual RGC types. Because we identified the individual RGC types, we were able to demonstrate that across all broader

functional groups, our modeling approach showed superior predictive performance using natural statistics than previous ones. This study showcased the use of our RGC type classifier and how it can be applied in such a study where a detailed cell type-resolved level is needed to demonstrate a general principle.

In summary, we developed an RGC type classifier that can reliably predict RGC types on recorded data using two-photon calcium imaging. The advantage of this classifier is that it only needs two distinct stimuli that are part of the standard repertoire used in most experiments; hence we can now even predict RGC types for previously recorded datasets. Additionally, the classifier shows an overall good prediction performance, which can be validated and improved compared to earlier approaches. Using this classifier, researchers can now perform cell type-resolved analyses on the level of the retinal output by simply using two short stimuli.

Removing Inter-experimental Variability

The first approach to analyze the effects of a neuromodulator on the retinal output signal was to investigate the overall response stability and consistency of RGC responses in the absence of any neuromodulator added to measure a baseline. Since neuromodulatory effects may be subtle, a careful control measurement must be employed. To this end, I recorded a dataset of paired RGC responses (Ctrl-dataset) under the same conditions as I would do with adding a neuromodulator to the system, i.e., each cell was recorded twice without any pharmacological manipulation.

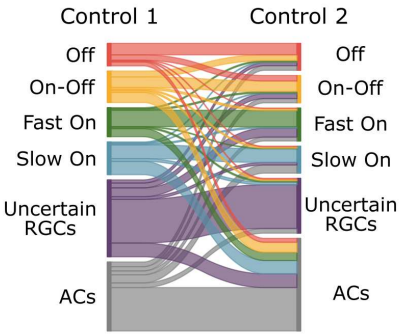

Figure 2: Application of the RGC type classifer on the first and second control recording. For each recorded RGC, the cell type was predicted using the RGC type classifier for both recordings, Control 1 and Control 2, separately. Sankey plot shows to which broader functional group each cell was assigned in both recording conditions.

Then, I applied the RGC type classifier on the first and second recordings of each cell separately. We hypothesized that cells in the Ctrl-dataset would be assigned to the same type, whereas in a complementary drug-dataset, certain types may be mapped to other types as distinct response features get modulated. This would allow us to track which RGC types were affected by the drug as well as to detect which features of a specific type would have been affected in particular. However, this approach did not seem to work as expected since RGC types in the Ctrl-dataset were not assigned to the same type consistently; hence we would not be able to separate those changes from drug-induced ones in the drug-dataset (Fig. 2). Such phenomena may be expected as related work by Zhao et al. (2020) showed response alterations due to external factors across recordings. This suggested that recording RGC populations twice without any pharmacological reagents may cause response changes and variations due to known and/or unknown experimental and biological factors, which may lead to unexpected and false-positive results of the data. These phenomena can be caused by external factors such as adaptation of the tissue, temperature fluctuations, experimental setup, and the experimenter. In our case, the main source of inter-experimental variability may be light adaptation throughout the experimental as all experiments were performed by the same experimenter and at the same setup. In general, there are several scenarios in which this may occur and a removal may be useful, e.g., pooling datasets and sequential recordings.

In this context, we aimed to address the issue of inter-experimental variability from two sides: *(i)* we developed a theoretical framework and showcased the removal of inter-experimental variability to demonstrate the importance of considering this issue when pooling datasets, and *(ii)* we established an experimental approach in which we leverage the knowledge about inter-experimental variability from a Ctrl-dataset onto a Drug-dataset to crystallize actual drug-induced effects. In the following section, I begin to elaborate on the theoretical approach, followed by the experimental one.

A Theoretical Approach to Remove Inter-experimental Variability

Nowadays, the process of pooling and merging datasets to acquire large-scale datasets for sophisticated analyses has become a standard procedure across disciplines and research fields (Baden et al., 2016; Franke et al., 2017). However, during this process of pooling data across several experiments in an iterative or sequential manner, various external biological or non-biological factors may introduce inter-experimental variability to the dataset, which can interfere with and distort the outcome of the analyses (Hicks et al., 2018; Zhao et al., 2020; Shah et al., 2022). In our cases, these factors of influence may originate from recording-related factors such as calcium buffering, temperature fluctuations, or photobleaching of the tissue.

As a consequence, we sought to address the issue of removing inter-experimental variability in a theoretical framework, yet a general approach to make it accessible and applicable for functional data across systems neuroscience. We showcased our theoretical framework on the example of merging two large-scale two-photon datasets of retinal BCs in response to various light stimuli. In brief, we addressed the problem of inter-experimental variability as one of domain shift and developed our theoretical framework `RAVE(+)` (**R**emoving, **A**dversarially, inter-experimental **V**ariability from **E**xperiments) in a setting of adversarial domain adaptation. We implemented two versions of `RAVE(+)` for different scenarios: *(i)* the unsupervised version (`RAVE`) can be used in cases where S is unknown and one simply wants to remove the inter-experimental variability between datasets, and *(ii)* the semi-supervised version (`RAVE+`) can be applied to scenarios where partial information about S is known for at least one of the datasets.

To remove inter-experimental variability, we implemented RAVE in a setting of an adversarial autoencoder (Fig. 3). Hereby, we want to learn a function f that maps the observation X, which is composed of the biological signal of interest S and the different domains D (different datasets or experiments), to a low-dimensional representation Z. This step represents the encoder of the autoencoder architecture. The function f has two competing objectives: *(i)* it should preserve as much information about S in Z as possible, and *(ii)* it should minimize information about D in Z at the same time. To do this, we need two additional functions. The function g performs a reconstruction of X to maximize the information about S in the learned lower representation Z, while the function h predicts D based on Z. Here, the function g represents the decoder of the autoencoder setting. The desired outcome is a lower representation Z learned from the function f from which we get a maximized reconstruction of S using the function g, while function h should only be able to predict D with chance-level. This would imply that we get a reconstruction of the signal without inter-experimental variability, as we cannot tell the domains apart.

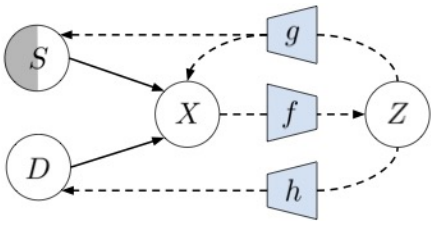

Figure 3: Schematic framework of `RAVE(+)`. Solid lines with arrows indicate a given relation. Dash lines with arrows represent learned and modelled relations. Functions are denoted with small letters and variables with capital ones. Adapted from Gonschorek et al. (2021).

To demonstrate the functionality of `RAVE(+)`, we used a pooled large-scale dataset of

light-induced responses of the 14 functional BCs types in the mouse retina (Euler et al., 2014). This pooled dataset consists of two published and openly-accessible datasets taken from two different studies (Franke et al., 2017; Szatko et al., 2020). In the following section, I will refer to the dataset from Franke et al., 2017 as *Dataset A* and Szatko et al., 2020 as *Dataset B*. To obtain those responses, both studies ubiquitously expressed the genetically encoded glutamate-sensing fluorescent reporter iGluSnFR (Marvin et al., 2013) in the mouse retina. Since the axon terminals of the distinct BC types stratify at different depths within the inner plexiform layer (IPL) of the inner retina, they imaged the light-induced responses of the glutamate output at those axon terminals. Notably, *Dataset A* identified and classified the 14 functional BC types using anatomically-guided functional clustering; however, this information was not available for *Dataset B*. In addition, even though both datasets used the same visual stimuli, i.e., local and full-field chirp stimuli, and expressed the same biosensor, they still exhibit slight differences, which complicates a simple merging of the two datasets.

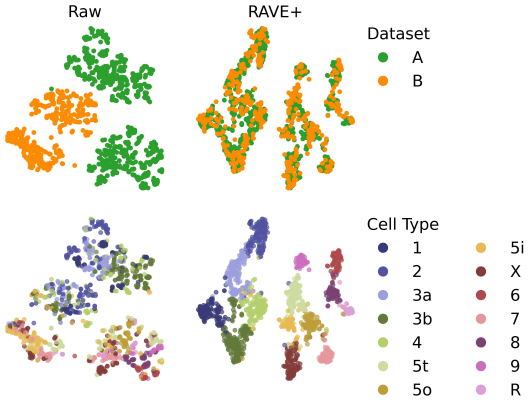

Figure 4: Embedding of the datasets. Low-dimensional t-SNE embeddings of the raw (left) and corrected (right) datasets. Color-code indicates datasets (top) and cell type labels (bottom). Adapted from Gonschorek et al. (2021).

Despite the similarity between the two datasets, we found clear evidence that inter-experimental variability has introduced a domain shift between them as both datasets are not able to be mixed, as seen by the low-dimensional embedding (Fig. 4; left). Here, the potential sources of inter-experimental variability originated from at least three external non-biological factors: (i) to acquire *Dataset B*, the authors used a slightly modified version of the chirp stimulus in which the stimulus section of the increasing frequency modulation is minimally slower compared to the one from *Dataset A*, (ii) the visual stimuli presented in *Dataset A* exhibited marginally different stimulus contrast because their visual stimulator did not use a gamma correction; thus there was no linearization of the intensity curve employed, and (iii) both datasets were imaged with different recording configurations, whereby *Dataset A* was generated using tangential scan recordings of distinct IPL depths and *Dataset B* by performing axial scan recordings and an electrically tunable lens enabling simultaneous recordings of the entire IPL. As a consequence, the problem can be formulated as the following: the observed responses of the 14 functional BCs in both datasets are represented as X and arise from the combination of the biological signal S and information about the domain D, whereas S is almost identical between *Dataset A* and *Dataset B*, yet D differs due to the external non-biological factors, resulting in inter-experimental variability.

RAVE+: Semi-supervised Removal of Inter-experimental Variability

For scenarios, in which there is only partial knowledge about the biological signal of interest available, e.g., cell type labels for only one dataset, yet is obscured by inter-experimental variability, we developed the semi-supervised version of RAVE (RAVE+). Here, RAVE+ aims to retain information about S, while minimizing D and predicting cell type labels, to enable an information transfer to the other dataset without containing inter-experimental variability.

As a first step to validate RAVE+, we simulated two large-scale BC datasets covering the 14 different cell types using the published BC model by Schröder et al. (2020a). To introduce inter-experimental variability between these datasets, we trained the model on the two versions of the chirp that were displayed to originally generate *Dataset A* and *Dataset B*. Then, we validated that the two datasets were affected by inter-experimental variability by projecting them into a low-dimensional t-SNE embedding (Hinton and Roweis, 2002), which showed that cells from both datasets were not overlapping (Fig 9b, Appendix, Study II). To test if we can transfer type information, we trained a cell type classifier on the responses of one dataset and then used this model to predict cell type labels for the other dataset. Since we have information about the ground-truth type labels of both simulated datasets, we could validate its prediction accuracy. Here, we found that the classifier only achieved a very poor prediction accuracy, indicating that the information transfer was strongly influenced by inter-experimental variability between the two datasets. However, when applying RAVE+ on the pooled dataset, we found that the cell type distributions overlapped when visualizing the datasets in a t-SNE embedding (Fig 9c, Appendix, Study II). In addition, by training the cell type classifier on the output embedding of RAVE+, we achieved a prediction performance of $\sim$0.99. When investigating the t-SNE embedding, color-coded according to dataset origin and cell type label, we found that the cell type labels form distinct islands containing cells from both datasets. Yet, when dealing with real experimental data, having ground-truth cell-type information available is usually not given. Thus, we needed another metric to validate the predicted cell type labels. To this end, we made use of the axon terminal stratification profile of the BC types as this represents a distinct cell type-specific anatomical feature.

To validate the predicted types per IPL depth, we used the electron-microscopy (EM) dataset comprising the probability distribution of the axon terminal stratification profiles of the 14 individual BC types (Helmstaedter, 2013; Behrens et al., 2016). We used this EM dataset as the expected cell type distribution across the IPL depth and compared this with the predicted type distribution of the raw data, i.e., without the removal, and RAVE+ output. Across all BC types, we found that the predicted type distributions using RAVE+, compared to the raw data, matched with the expected distributions from the EM dataset. We further quantified these results using various metrics. Consequently, we concluded that RAVE+ is indeed able to remove inter-experimental variability between the datasets and correctly predicts cell type labels in the simulated dataset (Fig 9d, Appendix, Study II).

After this validation, we applied RAVE+ to the real experimental data using our pooled large-scale dataset. Here, we also have partial information about the cell type information provided by *Dataset A* only. As for the simulated data, we now wanted to use *Dataset A* to predict cell type labels for *Dataset B* (Fig. 4; left). Similar to the simulated data, the t-SNE embedding of the raw datasets showed that they do not overlap, and when training the cell type classifier on *Dataset A* to predict type labels for *Dataset B*, we also found that this information was not able to be transferred (Fig. 4; left). Hence, this is a strong indicator of the presence of inter-experimental variability between these datasets. However, after the application of RAVE+, the low-dimensional embedding showed a clear mixture of *Dataset A* and *B*, as well as formed distinct cell type-specific islands (Fig. 4; right). As before, we leveraged

the IPL depth information by comparing the EM type distribution with the predicted ones from the raw dataset and RAVE+ output. As for the simulated data, RAVE+ showed clear improvements of type predictions compared to the raw data only (Fig. 5, Appendix, Study II); hence it allowed us to remove inter-experimental variability, while transferring type information, making it an appropriate framework for such scenarios.

RAVE: Unsupervised Removal of Inter-experimental Variability

In contrast to the supervised version of RAVE(+), we may not have partial information about S available for any of the datasets, i.e., no cell type labels. Such scenarios are common practice, for example, when several experiments with the same or similar conditions are conducted and pooled as large-scale dataset for further analyses. Here, the pooling process may introduce inter-experimental variability, which may influence the overall outcome and interpretation of downstream analyses. To this end, we developed the unsupervised version of RAVE, in which the framework provides a reconstruction of the functional traces without inter-experimental variability to simplify the merging process of such datasets.

To showcase the feasibility of such a downstream analysis after the removal of the inter-experimental variability between *Dataset A* and *Dataset B*, we aimed to reproduce a specific analysis on *Dataset B* that was previously performed on *Dataset A*. Here, the analysis revealed a fundamental biological feature of BCs: visual full-field stimulation causes a decorrelation of BC type responses in comparison to local stimulation as the inhibitory surround feedback, driven by ACs, is activated (Franke et al., 2017). This has the consequence that BC types sharing the same response polarity have high response correlations to a local stimulus, while exhibiting a decorrelation to the full-field one. As this feature is essential for BC types, we would expect to find this in *Dataset B* as well, since they share the underlying biological signal. However, when training a classifier on the raw data to predict cell type labels for *Dataset B* and then replicating the analysis, we did not find the expected feature to be presented, and therefore failed to replicate the results (Fig. 6, Appendix, Study II). As described and showed above, the subtle domain differences between the two datasets caused the presence of inter-experimental variability; thus making it difficult to transfer information from *Dataset A* to *Dataset B*. Yet, after the application of RAVE on the raw traces, we trained the classifier on the reconstructed traces and repeated the analysis, revealing the existence of the fundamental feature in *Dataset B* as expected.

Showcasing this, we were able to demonstrate the usage of RAVE in scenarios without additional information, such as cell type labels. Due to the successful removal of inter-experimental variability from functional traces, pooling large-scale datasets in systems neuroscience may produce more reliable results in the future as the domain shift can be minimized.

An Experimental Approach: Paired Control Recordings

Addressing the problem of inter-experimental variability in a theoretical manner may be useful for datasets that were already recorded and now need to be pooled for further analyses. Nevertheless, it would be advantageous to minimize and control such factors of influence as far as possible already during the data acquisition process. In the aforementioned showcased examples using RAVE(+), we revealed the existence and associated problems of inter-experimental variability, which causes large discrepancies between datasets containing the same underlying biological signal; thus making it difficult to reliably merge them. However, studying drug-induced effects adds an extra layer of complexity as the

expected signal modulation may be diverse and differential depending on the cell types and the affected neural circuits. Therefore, tackling this issue experimentally may help to minimize such factors as well as to distill drug-induced effects separately from potential other factors. As a consequence, I proposed and validated a recording protocol to account for biological/non-biological external factors and transfer those to drug-induced experiments.

To this end, I generated two datasets of sequentially recorded light-induced RGC responses, resulting in paired RGC responses: *(i)* a Ctrl-dataset in which RGC populations were recorded twice without any pharmacological perturbation (Ctrl 1 & Ctrl 2), and *(ii)* a Drug-dataset recorded in the same manner, but in which the drug is perfused to the *ex vivo* retina before the second recording (Ctrl 1 & Drug) (Fig. 1, Appendix, Study III). The Ctrl-dataset served the purpose of investigating if there are systematic and/or non-systematic changes across RGC types, i.e., if all, only certain, or none of the RGC types show signal modulations as a consequence of the sequential recordings. Then, the Drug-dataset was used to study the drug-induced effects of the used pharmacological reagent. Notably, the Drug-dataset also contains the potential systematic and/or non-systematic modulations in addition to the drug-induced ones.

To identify functional types and study how each of these is affected, I applied the RGC type classifier to the light-induced responses of the Ctrl 1 recordings, as these represent the baseline activity without any additional manipulations. As described earlier, applying the RGC type classifier on Ctrl 1 and Ctrl 2 of the Ctrl-dataset resulted in a discrepancy of mapping types. To account for such potential time-dependent effects, I first computed the differences between the Ctrl 1 and Ctrl 2 recordings of each cell. This provided a baseline metric of non-drug-induced effects. Then, the same procedure was applied to the Drug-dataset. Next, I computed the difference of the difference of the Ctrl- and Drug-dataset. Since we would expect similar time-dependent effects for the same RGC types in both datasets, computing this difference would eliminate the time-dependent effects from the Drug-dataset and reveal only the actual drug-induced ones.

Usually, one may claim that this kind of recording procedure may be redundant since one could simply perform a wash-out recording after the drug application. However, this was not feasible in our recording setting as a third recording may result in a overall reduction of the signal-to-noise ratio; thus reliable recordings and conclusions may be difficult to make. Technically, this procedure can be applied to most functional recording experiments where one can perform sequential recordings of the same neurons, such as patch-clamp or MEA recordings. In the case of MEA, the procedure might be more challenging to apply as we do not know if we deal with the same cells over time. As long as the detected waveforms of the single neurons remain similar, this does not constitute a big hurdle. But if the application of a neuromodulator, or any other drug, may influence the shape of the waveform, then the spike sorting to separate neurons may cause the separation of neurons not by their location on the array but rather by condition (e.g., Ctrl 1 & Drug) (Buccino et al., 2022). Also, those time-dependent effects may be a consequence of the *ex vivo* condition of the tissue. It would be interesting to test if neurons display such time-dependent changes in the *in vivo* setting.

In summary, I proposed a recording protocol to account for time-dependent effects by recording a Ctrl-dataset that grasps cell type-specific effects, which can be transferred to a Drug-dataset recorded in the same manner. This allows us to untangle such effects from drug-induced ones. In the next section, I describe and discuss the results of this procedure on a showcased example of how the neuromodulator nitric oxide (NO) affects the retinal output, how this recording protocol revealed type-specific time-dependent effects, as well as why this may be useful for future research on neuromodulatory effects and even other pharmacological reagents.

Case Study: Nitric Oxide and its Effects on the Retinal Output

I began this thesis by highlighting the importance of studying neuromodulatory effects on neural circuits and that we need to investigate them in detail on a systematic level. The so far presented sections provided the tools to study those effects on the retinal output signal. Therefore, for the final part of my thesis, I bring together the individual parts of my research in one case study in which I elaborate on how I use these tools to study the effects of the neuromodulator NO and its implications on the signal of the diversity of RGC types.

Nitric Oxide: A Double-edged Sword

Nitric oxide (NO) is a gaseous signaling molecule, also termed a gasotransmitter, that exhibits a broad spectrum of physiological functions across various biological systems (Snyder and Bredt, 1992; Bruckdorfer, 2005). The most prominent example of its physiological function is its action as vasodilator to cause a dilation of smooth muscle cells in blood vessels (Knowles and Moncada, 1992). However, NO is also vastly found across the central and peripheral nervous system, where it mainly acts as a neuromodulator to alter the activity of neural circuits (Bredt and Snyder, 1990; Garthwaite, 2008). To synthesize NO, it requires the enzymatic conversion of L-arginine by nitric oxide synthase (NOS) into NO and L-citrulline (Alderton et al., 2001; Daff, 2010). For this process, additional cofactors such as NADPH, oxygen and calcium (depending on the isoform) are required to catalyze this reaction. Overall, there are three different isoforms of NOS: *(i)* calcium-dependent neuronal NOS (nNOS), which is mainly expressed in neurons, *(ii)* inducible NOS (iNOS) expressed during pathological and inflammatory states, and *(iii)* endothelial NOS (eNOS) mostly expressed in the vascular endothelium such as blood vessels.

Within the retina, nNOS constitutes the dominant isoform in the context of visual signal processing to produce NO (Dawson et al., 1991; Tsumamoto et al., 2002). The production of NO via nNOS is triggered by the calcium influx into the cell under the regulation of calmodulin. The enzyme nNOS is present in various retinal cell classes, although in mice, nNOS is mainly found in the inner retina (Vielma et al., 2012). In particular, nNOS is primarily expressed in a subgroup of ACs (Dawson et al., 1991), namely nNOS1-ACs and nNOS2-ACs, whereby it was shown that nNOS2-ACs are the dominant source of NO in a light-dependent manner (Jacoby et al., 2018).

NO has two modes of actions to modulate neural activity: *(i)* NO can activate soluble guanylate cyclase (sGC/NO-GC) and thereby trigger an intracellular signaling cascade increasing the concentration of the second messenger cGMP, which in turn will cause downstream modulations of several target proteins (Garthwaite, 2005), and *(ii)* S-nitrosylation (cGMP-independent), in which NO directly binds to and nitrosylates specific receptors, ion channels and proteins to modulate, for example, neurotransmitter release (Ahern et al., 2002; Hess et al., 2005; Knott and Bossy-Wetzel, 2009). From these two signaling pathways, the former one is thought to represent the main mode of how NO operates within the retina.

To date, there is no comprehensive view of the functionality and effects of NO on the retinal signal processing, even though the signaling pathway was studied in depth (e.g., Blom et al., 2012). Like the neuromodulator dopamine, the effects of NO are essentially associated with mechanisms of light adaptation, whereby it was discussed why the retina needs two light adaptational systems (Jacoby et al., 2018). Researcher hypothesized that dopamine is responsible for light adaptation on the time-scale of the circadian rhythm, i.e., throughout the day, whereas NO rather facilitates light adaptation over a short time-scale by affecting and

modulating the signal transmission between the scotopic and photopic pathways (Nir et al., 2000). It was shown that NO can regulate these pathways using several mechanisms such as decoupling of gap junctions between On-CBs and the AII, which is the intersection of the rod- and cone-pathway (Mills and Massey, 1995). In addition, NO modulates the temporal response properties of CBCs in scotopic conditions by altering the response gain of Off-CBCs (Vielma et al., 2014) and increasing the one of On-CBCs to extend their operational response spectrum (Nawy and Jahr, 1990; Snellman and Nawy, 2004).

In comparison, studies on the effects of NO to modulate the retinal output are rare and mainly focus on a handful of cells in a more cell type-unspecific way by investigating On- and Off-RGCs. Here, it was found that NO may decrease the response strength of On- and Off types, yet not resolved on type-level (Wang et al., 2003).

As a consequence, I sought to tackle this issue and describe a case study in the next section, in which I analyze NO-induced effects from a global perspective using a population-based and cell type-resolved approach.

Nitric Oxide Mediates Differential Effects in Mouse Retinal Ganglion Cells

So far, the majority of research studying the effects of neuromodulators in the retina mainly focused on the localization of producing enzymes, receptors, and circuit mechanisms. However, studying the functional implications of neuromodulators on the retinal output, i.e., on RGC level, are rare. A recent publication studied the effects of dopamine on RGC type-level (Warwick et al., 2023), and other research groups mainly focused on analyzing neuromodulatory effects on single-cell level (Wang et al., 2003), however no study investigated the effects of NO on population level, yet. Since RGCs encode the visual world in a diversity of parallel feature channels represented by various types and not as a uniform signal, resolving the retinal output signal down to cell type-level is key to understand early vision. Due to this vast diversity of RGC types, we hypothesized that NO may have differential effects across functional types.

To test this hypothesis, I employed two-photon calcium imaging of RGC populations (Baden et al., 2016; Euler et al., 2019). For this, I expressed the synthetic fluorescent calcium indicator Oregon Green-BAPTA 1 in the GCL using bulk electroporation (Briggman and Euler, 2011) and displayed various light stimuli. Here, I mainly used the full-field chirp and moving bar stimuli as they were needed for the functional classification using the RGC type classifier. To account of potential time-dependent effects, I employed the earlier described paired recording protocol, in which I recorded two separated paired datasets: (i) Ctrl-dataset (Ctrl 1 & Ctrl 2) without pharmacological manipulation, and (ii) NO-dataset (Ctrl & NO), whereby DETA/NO (NO-donor) was added before the second recording (Fig. 1, Appendix, Study III). For the functional identification of the different types, I applied the previously described RGC type classifier exclusively on the responses to the chirp and moving bar stimulus of the first recordings of both datasets (Ctrl 1, Fig. 2, Appendix, Study III). This was necessary as these responses represent the baseline activity of the cell types and are neither influenced by time-dependent nor NO-induced effects (depending on the dataset).

After the cell type classification, the first step was to investigate the response stability between two recordings in the absence of pharmacological perturbation using the Ctrl-dataset. To account for systematic and/or specific time-dependent changes, I employed the paired recording protocol as described earlier. While computing the type-specific differences between the paired responses of the Ctrl-dataset, I found that the majority of types across all five larger RGC groups (Off, On-Off, Fast On, Slow On & Uncertain RGCs) exhibited response

modulations, which we termed *adaptational cell type-specific changes* (Fig. 3, Appendix, Study III).

Next, I repeated this analysis for the NO-dataset and compared the results with the findings of the Ctrl-dataset. Surprisingly, almost all types that displayed such type-specific changes in the Ctrl-dataset, i.e., an increase or decrease in their activity during the second recording, also showed the same trend in the NO-dataset. Looking at these results, I saw that those effects were actually not induced by NO per se. Even if NO did not exhibit a broad, systematic effect across RGC types, this finding constitutes an important step towards in studying such drugs as we can now identify apparent drug-induced effects from our data.

However, even if almost all types did not show NO-induced effects, there was still one type that exhibited stable responses during the control recordings, yet displayed a clear activity increase during the NO-perfused condition. This type is part of the subgroup 'Uncertain RGCs' and is termed 'Off suppressed-by-contrast 2' (G_{32}) (Baden et al., 2016). Its characteristic feature is a high baseline activity that gets suppressed in the presence of light onsets. To analyze the adaptational cell type-specific changes and the NO-induced effects on a more detailed level, I decomposed the responses to the chirp and moving bar into 8 features (six features for the chirp and two for the moving bar) (Fig. 4, Appendix, Study III). By computing the feature-wise difference between types for the Ctrl- and NO-datasets, I found that these adaptational cell type-specific changes were very accurate and replicable. To eliminate, and thus disentangle the influence of these type-specific changes from the NO-induced effects, I subtracted the type-wise differences of the Ctrl-dataset from the differences of the NO-dataset. As a consequence, only NO-induced effects may remain as the time-dependent changes were supposed to be the same across datasets. Overall, I found that G_{32} was still the only candidate exhibiting NO-induced effects across all relevant features. Interestingly, from Baden et al. (2016), we could speculate that G_{32} may comprise more than just one type, but three subtypes due to its coverage factor of >2. As a consequence, I pooled the chirp and moving bar responses of RGCs classified as G_{32} of the Ctrl 1 recordings from both the Drug- and Ctrl-dataset pairs and performed a functional clustering (Fig. 5, Appendix, Study III). To this end, I extracted the visual features using sparse principal component analysis (Zou et al., 2006) and then performed Mixture of Gaussian (Reynolds et al., 2009) on the standardized features. By testing several models with different numbers of clusters, initializing the process multiple times and computing the average Bayesian Information Criterion (BIC) for each model (Neath and Cavanaugh, 2012), the model with the lowest BIC was selected. The clustering resulted in three distinct clusters for G_{32}, matching with the proposed number of clusters in Baden et al. (2016). Next, I compared each cluster with the Ctrl 2 and DETA/NO condition separately. Here, I found for the Ctrl-dataset that the three clusters showed stable responses across recordings, whereas in the NO-dataset the clusters mostly lost their characteristic feature of being suppressed-by-contrast. However, all three clusters, likely being three RGC subtypes based on the coverage factor of G_{32}, exhibited different levels of excitation, indicating a more differential effect of NO on this subgroup of RGCs.

This finding was surprising as NO is vastly prominent in the inner retina, yet only one distinct type, and its subtypes, is modulated in its response to temporal stimuli. Next, I sought to investigate if NO may modulate RGC types in their spatial domain by recording their responses to binary dense noise to compute their spatial receptive fields (sRFs) using spike-triggered averages (STA) (Chichilnisky, 2001) (Fig. 6, Appendix, Study III). Interestingly, when comparing sRF diameter and surround index differences between Ctrl and NO conditions of paired RGC types, I found that they were highly stable across RGC types. Consequently, this may imply that NO has a distinct modulatory effect on a subgroup of RGCs in their temporal response domain.

In summary, I systematically investigated the effects of NO on the RGC signal processing on cell type-level. For this, we used the RGC type classifier we developed for a previous study and the paired control recording protocol. The classifier enabled us to resolve those effects on cell type-level, whereas the recording protocol helped us to identify time-dependent effects and then disentangle those effects from NO-induced ones. This experimental pipeline can now be applied to any kind of neuromodulator in the context of the retinal output signal.

Discussion

In this thesis, the overall aim was to establish an experimental and computational pipeline to uncover the neuromodulatory effects of NO on the retinal output signal. Neuromodulators have major influences on neural circuits dynamics, yet with common tools, their contribution and implications are often neglected, which may cause knowledge gaps when studying such systems on a functional level. Therefore, the here presented tools and findings shed light on the importance of their function and potential pitfalls when studying their pathways and interactions within a system.

In Study I, we developed an RGC type classifier that enables a fast and reliable identification of functional RGC types recorded using two-photon calcium imaging and two light stimuli. In the context of studying neuromodulatory effects in the retina, cell type-resolved analyses provide a crucial aspect due to the potentially specific and diverse functionality of neuromodulators on the RGC type diversity. To date, other studies that investigate effects on RGC type population level either performed clustering approaches followed by matching the individual clusters to known types using references such as `rgctypes.org` (Goetz et al., 2022; Goldin et al., 2022; Warwick et al., 2023) or identified RGC types based on their receptive field characteristics (Shi et al., 2019). Yet, in most cases, using a more standardized procedure such as a RGC type classifier as ours would simplify and unify functional RGC typing in a way that researchers in the field would be able to identify the same types across experimental approaches.

In Study II, and partly Study III, we addressed the problem of pooling datasets and sequential recordings of the same cells. Because we studied the effect of a particular neuromodulator, one has to perform control measurement as baseline and then add the neuromodulator of interest to observe its effect on the system. Therefore, sequential recordings and pooling of data are necessary proceedings to conduct. Usually, the majority of studies either perform additional wash-out recordings after the drug-perfused one or simply record the drug-condition without a wash-out, assuming the observed behavior being exclusively induced by the drug. However, our paired recording protocol clearly showed that, at least under our experimental conditions, such assumptions would lead to apparent drug-induced effects for the majority of RGC types, i.e., effects that are not caused by the drug of interest, but the experimental conditions and adapational effects. Our first approach to remove such inter-experimental variabilities from a pooled dataset in a theoretical framework using `RAVE(+)` outperformed existing algorithms. The framework adds important value to the field of systems neuroscience as researchers can merge datasets and potentially detect more subtle effects that may be obscured by inter-experimental variability otherwise. Yet, its application may be limited to datasets with the same underlying biological signal; hence we were not able to apply `RAVE(+)` in its current setting onto the NO-dataset. The biggest drawback here is that the underlying biological signal of interest, i.e., light-evoked responses, between Ctrl 1 and NO may not be the same; hence `RAVE(+)` may remove actual drug-induced effects by trying to match the datasets. Nevertheless, adapting `RAVE(+)` in the way that it learns the mapping between Ctrl 1 and Ctrl 2 of the Ctrl-dataset may support transferring information onto Ctrl 1 and NO to allow a removal of the underlying inter-experimental variability without the removal of the drug-induced effects. This approach is partly similar to the paired recording protocol, yet it does not constitute a linear procedure and could therefore highlight even more subtle differences. Another approach to resolve drug-induced effects on a continuous time axis in such a setting would make use of Gaussian processes (GP). In their study, Rogerson et al. (2019) used this approach to compute the model difference of responses between two similar stimuli to estimate if and where these signals are different. This is analogous to our recording

protocol and may enable a precise estimate of which features are affected by time-dependent and drug-induced effects in a non-linear manner. However, to properly compute GPs, the sampling rate of the responses need to be higher than the one used in our study, otherwise it is difficult to estimate the uncertainty in between the single recording points. To solve this issue, we could employ a different recording configuration as suggested by Rogerson et al. (2019). However, since we wanted to make use of the RGC type classifier, we had to record the data in this particular configuration; thus a GP analysis was therefore not easily possible. Nevertheless, we were able to develop and validate necessary tools to address several keystone aspects to study neuromodulatory effects in the retina.

In the final part of my thesis, I summarized the results of Study III, in which we put the previously developed methods together to study the functional effects of the neuromodulator NO on the retinal output at cell type-resolution, while accounting for time-dependent effects. In our study, we had two main findings: *(i)* one third of the RGC types showed cell type-specific temporal adaptations already without pharmacological manipulation, and *(ii)* NO seemed to affect a distinct group of RGC types in a differential manner. The former finding was discussed above, however, the later one was surprising to us as NO and its receptors are vastly expressed and released throughout the inner retina (Vielma et al., 2014), yet mainly affecting the functionality of one particular subgroup. Interestingly, the neuromodulator dopamine was shown to act via distinct signaling pathways to differentially modulate the receptive field structures of specific RGC types (Warwick et al., 2023). Even though they found differential effects across specific types, not all RGC types were affected by dopamine, indicating type-specific effects, which is in line with our findings on NO. Indeed, we found NO induced functional modulation operating on a cell type-specific level, yet the underlying mechanisms remain to be investigated further. Due to the two parallel pathways of NO (sGC-cGMP and S-nitrosylation), further experiments are needed to disentangle those pathways. Here, two experimental approaches may be suitable. To investigate if NO functions via the cGMP-dependent pathway, one may block the sGC/NO-GC receptors using ODQ and then add the NO-donor to the system. If the previously observed effects still occur, one could speculate that the neuromodulatory effects are signaled through S-nitrosylation, the cGMP-independent pathway. Otherwise, the likelihood for the cGMP-dependent pathway as the main signaling pathway is increased. To verify this, dual-imaging of cGMP and calcium may be useful. To this end, one could use the specific transgenic mice line cGi500, which ubiquitously expresses a FRET-based cGMP-sensor (Thunemann et al., 2013), and combines it with bulk electroporation of a calcium indicator (Briggman and Euler, 2011). The calcium imaging would allow the identification of different RGC types using the RGC type classifier as well as to correlate calcium signals with cGMP signal modulations. First preliminary experiments and data suggest a differential and cell type-specific modulation of cGMP within the GCL.

An additional aspect to consider is the concentration of the NO-donor, DETA/NO, which in our study, at the first glance, lays outside the proposed endogenous physiological range (Eldred and Blute, 2005). Starting such an approach with higher concentrations in this experimental setting is useful as we had to overcome several barriers to highlight potential effects. Moreover, one needs to take into account that NO, even though bound as DETA/NO exhibiting a half-life time of $\sim$24 hours at normal physiological temperature (Hanson et al., 1995), is a gas; thus the actual concentration locally in the tissue is likely much lower. Moreover, since we used a whole-mount preparation of the *ex vivo* intact retina, NO may take more time to diffuse into the tissue compared to a vertical slice preparation, which may affect the final concentration as well, yet it mimics the natural conditions by keeping the entire retinal network intact.

To map the individual functional RGC types, we had to make use of two artificial stimuli, i.e., chirp and moving bar. In general, the usage of artificial stimuli provides several advantages such as a standardized stimulus feature space and fairly straightforward analyses. However, their stimulus space is limited and could be extended by using natural stimuli to approximate more natural conditions in which neuromodulators are released and function (Qiu et al., 2021). A recent study by Höfling et al. (2022) used a joint approach, in which they trained a neural network on RGC responses to natural movies. Here, they showed that a distinct RGC type is mainly responsible for the visual transition from ground scenes to sky scenes. Using artificial stimuli, this finding would probably not have been discovered in this way. To this end, such an approach seems promising and should be employed when studying neuromodulators in the retina.

In summary, in this thesis, I aimed to shine light on the importance of studying neuromodulators in the context of visual signal processing. Neuromodulators certainly add complexity to the overall functionality of neural circuits that may be neglected or simplified otherwise. To this end, I demonstrated how a joint experimental and computational approach enables a precise analysis of their functional implications for vision considering several experimental pitfalls. Especially, causing awareness of potential inter-experimental variability when merging datasets as well as adaptational effects in such pharmacological studies should contribute to the field in a way that researchers may be careful when interpreting such data. Further advances in experimental and computational techniques will broaden the picture of neuromodulators in the future.

References

Ahern, Gerard P, Vitaly A Klyachko, and Meyer B Jackson (2002). "cGMP and S-nitrosylation: two routes for modulation of neuronal excitability by NO". In: *Trends in neurosciences* 25.10, pp. 510–517.

Alderton, Wendy K, Chris E Cooper, and Richard G Knowles (2001). "Nitric oxide synthases: structure, function and inhibition". In: *Biochemical journal* 357.3, pp. 593–615.

Alger, Bradley E (2002). "Retrograde signaling in the regulation of synaptic transmission: focus on endocannabinoids". In: *Progress in neurobiology* 68.4, pp. 247–286.

Applebury, ML, MP Antoch, LC Baxter, LLY Chun, JD Falk, F Farhangfar, K Kage, MG Krzystolik, LA Lyass, and JT Robbins (2000). "The murine cone photoreceptor: a single cone type expresses both S and M opsins with retinal spatial patterning". In: *Neuron* 27.3, pp. 513–523.

Asari, Hiroki and Markus Meister (2012). "Divergence of visual channels in the inner retina". In: *Nature neuroscience* 15.11, pp. 1581–1589.

Baden, Tom, Philipp Berens, Matthias Bethge, and Thomas Euler (2013). "Spikes in mammalian bipolar cells support temporal layering of the inner retina". In: *Current Biology* 23.1, pp. 48–52.

Baden, Tom, Philipp Berens, Katrin Franke, Miroslav Román Rosón, Matthias Bethge, and Thomas Euler (2016). "The functional diversity of retinal ganglion cells in the mouse". In: *Nature* 529.7586, pp. 345–350.

Bae, J Alexander, Shang Mu, Jinseop S Kim, Nicholas L Turner, Ignacio Tartavull, Nico Kemnitz, Chris S Jordan, Alex D Norton, William M Silversmith, Rachel Prentki, et al. (2018). "Digital museum of retinal ganglion cells with dense anatomy and physiology". In: *Cell* 173.5, pp. 1293–1306.

Bargmann, Cornelia I (2012). "Beyond the connectome: how neuromodulators shape neural circuits". In: *Bioessays* 34.6, pp. 458–465.

Barlow, HB, Roo Fitzhugh, and SW Kuffler (1957). "Change of organization in the receptive fields of the cat's retina during dark adaptation". In: *The Journal of physiology* 137.3, p. 338.

Baylor, Denis A (1987). "Photoreceptor signals and vision. Proctor lecture." In: *Investigative ophthalmology & visual science* 28.1, pp. 34–49.

Behrens, Christian, Timm Schubert, Silke Haverkamp, Thomas Euler, and Philipp Berens (2016). "Connectivity map of bipolar cells and photoreceptors in the mouse retina". In: *Elife* 5, e20041.

Behrens, Christian, Shubhash Chandra Yadav, Maria M Korympidou, Yue Zhang, Silke Haverkamp, Stephan Irsen, Anna Schaedler, Xiaoyu Lu, Zhuohe Liu, Jan Lause, et al. (2022). "Retinal horizontal cells use different synaptic sites for global feedforward and local feedback signaling". In: *Current Biology* 32.3, pp. 545–558.

Berry, Michael H, Michael Moldavan, Tavita Garrett, Marc Meadows, Olga Cravetchi, Elizabeth White, Henrique von Gersdorff, Kevin M Wright, Charles Allen, and Benjamin Sivyer (2022). "A subtype of melanopsin ganglion cells encodes ground luminance". In: *bioRxiv*, pp. 2022–04.

Berson, DM, FA Dunn, and M Takao (2001). "Phototransduction by ganglion cells innervating the circadian pacemaker." In: *Investigative Ophthalmology & Visual Science*. Vol. 42. 4. Assoc Research Vision Ophthalmology Inc 9650 Rockville Pike, Bethesda, MD ..., S113–S113.

Blom, Jan, Tom Giove, Monika Deshpande, and William D Eldred (2012). "Characterization of nitric oxide signaling pathways in the mouse retina". In: *Journal of Comparative Neurology* 520.18, pp. 4204–4217.

Bredt, David S and Solomon H Snyder (1990). "Isolation of nitric oxide synthetase, a calmodulin-requiring enzyme." In: *Proceedings of the National Academy of Sciences* 87.2, pp. 682–685.

Breuninger, Tobias, Christian Puller, Silke Haverkamp, and Thomas Euler (2011). "Chromatic bipolar cell pathways in the mouse retina". In: *Journal of Neuroscience* 31.17, pp. 6504–6517.

Briggman, Kevin L and Thomas Euler (2011). "Bulk electroporation and population calcium imaging in the adult mammalian retina". In: *Journal of neurophysiology* 105.5, pp. 2601–2609.

Bruckdorfer, Richard (2005). "The basics about nitric oxide". In: *Molecular aspects of medicine* 26.1-2, pp. 3–31.

Buccino, Alessio P, Samuel Garcia, and Pierre Yger (2022). "Spike sorting: new trends and challenges of the era of high-density probes". In: *Progress in Biomedical Engineering* 4.2, p. 022005.

Cao, Luxiang, Todd A Blute, and William D Eldred (2000). "Localization of heme oxygenase-2 and modulation of cGMP levels by carbon monoxide and/or nitric oxide in the retina". In: *Visual Neuroscience* 17.3, pp. 319–329.

Chen, CK (2005). "The vertebrate phototransduction cascade: amplification and termination mechanisms". In.

Chen, S-K, TC Badea, and S Hattar (2011). "Photoentrainment and pupillary light reflex are mediated by distinct populations of ipRGCs". In: *Nature* 476.7358, pp. 92–95.

Chichilnisky, EJ (2001). "A simple white noise analysis of neuronal light responses". In: *Network: computation in neural systems* 12.2, p. 199.

Cleland, BG, MW Dubin, and WR Levick (1971). "Sustained and transient neurones in the cat's retina and lateral geniculate nucleus". In: *The Journal of physiology* 217.2, pp. 473–496.

Dacey, Dennis (2004). "20 origins of perception: Retinal ganglion cell diversity and the creation of parallel visual pathways". In: *The Cognitive Neurosciences Iii.*, p. 281.

Dacey, Dennis M (1990). "The dopaminergic amacrine cell". In: *Journal of Comparative Neurology* 301.3, pp. 461–489.

Daff, Simon (2010). "NO synthase: structures and mechanisms". In: *Nitric oxide* 23.1, pp. 1–11.

Dawson, Ted M, David S Bredt, Majid Fotuhi, Paul M Hwang, and Solomon H Snyder (1991). "Nitric oxide synthase and neuronal NADPH diaphorase are identical in brain and peripheral tissues." In: *Proceedings of the National Academy of Sciences* 88.17, pp. 7797–7801.

Demb, Jonathan B and Joshua H Singer (2012). "Intrinsic properties and functional circuitry of the AII amacrine cell". In: *Visual neuroscience* 29.1, pp. 51–60.

DeVries, Steven H (2000). "Bipolar cells use kainate and AMPA receptors to filter visual information into separate channels". In: *Neuron* 28.3, pp. 847–856.

Dhande, Onkar S and Andrew D Huberman (2014). "Retinal ganglion cell maps in the brain: implications for visual processing". In: *Current opinion in neurobiology* 24, pp. 133–142.

Diamond, Jeffrey S (2017). "Inhibitory interneurons in the retina: types, circuitry, and function". In: *Annual review of vision science* 3, pp. 1–24.

Dong, Dawei W and Joseph J Atick (1995). "Statistics of natural time-varying images". In: *Network: Computation in Neural Systems* 6.3, p. 345.

Drinnenberg, Antonia, Felix Franke, Rei K Morikawa, Josephine Jüttner, Daniel Hillier, Peter Hantz, Andreas Hierlemann, Rava Azeredo da Silveira, and Botond Roska (2018). "How diverse retinal functions arise from feedback at the first visual synapse". In: *Neuron* 99.1, pp. 117–134.

Eldred, William D and Todd A Blute (2005). "Imaging of nitric oxide in the retina". In: *Vision research* 45.28, pp. 3469–3486.

Euler, Thomas, Peter B Detwiler, and Winfried Denk (2002). "Directionally selective calcium signals in dendrites of starburst amacrine cells". In: *Nature* 418.6900, pp. 845–852.

Euler, Thomas, Katrin Franke, and Tom Baden (2019). "Studying a light sensor with light: multiphoton imaging in the retina". In: *Multiphoton Microscopy*, pp. 225–250.

Euler, Thomas, Silke Haverkamp, Timm Schubert, and Tom Baden (2014). "Retinal bipolar cells: elementary building blocks of vision". In: *Nature Reviews Neuroscience* 15.8, pp. 507–519.

Franke, Katrin, Philipp Berens, Timm Schubert, Matthias Bethge, Thomas Euler, and Tom Baden (2017). "Inhibition decorrelates visual feature representations in the inner retina". In: *Nature* 542.7642, pp. 439–444.

Freedman, Melanie S, Robert J Lucas, Bobby Soni, Malcolm Von Schantz, Marta Muñoz, Zoë David-Gray, and Russell Foster (1999). "Regulation of mammalian circadian behavior by non-rod, non-cone, ocular photoreceptors". In: *Science* 284.5413, pp. 502–504.

Garthwaite, John (2005). "Dynamics of cellular NO-cGMP signaling". In: *Front Biosci* 10, pp. 1868–1880.

Garthwaite, John (2008). "Concepts of neural nitric oxide-mediated transmission". In: *European Journal of Neuroscience* 27.11, pp. 2783–2802.

Gastinger, Matthew J, Alistair J Barber, Noga Vardi, and David W Marshak (2006a). "Histamine receptors in mammalian retinas". In: *Journal of Comparative Neurology* 495.6, pp. 658–667.

Gastinger, Matthew J, Ning Tian, Tamas Horvath, and David W Marshak (2006b). "Retinopetal axons in mammals: emphasis on histamine and serotonin". In: *Current eye research* 31.7-8, pp. 655–667.

Ghosh, Krishna K, Sascha Bujan, Silke Haverkamp, Andreas Feigenspan, and Heinz Wässle (2004). "Types of bipolar cells in the mouse retina". In: *Journal of Comparative Neurology* 469.1, pp. 70–82.

Goetz, Jillian, Zachary F Jessen, Anne Jacobi, Adam Mani, Sam Cooler, Devon Greer, Sabah Kadri, Jeremy Segal, Karthik Shekhar, Joshua R Sanes, et al. (2022). "Unified classification of mouse retinal ganglion cells using function, morphology, and gene expression". In: *Cell reports* 40.2, p. 111040.

Goldin, Matías A, Baptiste Lefebvre, Samuele Virgili, Mathieu Kim Pham Van Cang, Alexander Ecker, Thierry Mora, Ulisse Ferrari, and Olivier Marre (2022). "Context-dependent selectivity to natural images in the retina". In: *Nature Communications* 13.1, p. 5556.

Goldstein, Ira M, Philipp Ostwald, and Steven Roth (1996). "Nitric oxide: a review of its role in retinal function and disease". In: *Vision research* 36.18, pp. 2979–2994.

Gollisch, Tim and Markus Meister (2010). "Eye smarter than scientists believed: neural computations in circuits of the retina". In: *Neuron* 65.2, pp. 150–164.

Gonschorek, Dominic, Larissa Höfling, Klaudia P Szatko, Katrin Franke, Timm Schubert, Benjamin Dunn, Philipp Berens, David Klindt, and Thomas Euler (2021). "Removing inter-experimental variability from functional data in systems neuroscience". In: *Advances in Neural Information Processing Systems* 34, pp. 3706–3719.

Grumet, Andrew E, John L Wyatt Jr, and Joseph F Rizzo III (2000). "Multi-electrode stimulation and recording in the isolated retina". In: *Journal of neuroscience methods* 101.1, pp. 31–42.

Hanson, Stephen R, Thomas C Hutsell, Larry K Keefer, Daniel L Mooradian, and Daniel J Smith (1995). "Nitric oxide donors: a continuing opportunity in drug design". In: *Advances in Pharmacology* 34, pp. 383–398.

Harvey, John D and Thomas Heinbockel (2018). "Neuromodulation of synaptic transmission in the main olfactory bulb". In: *International Journal of Environmental Research and Public Health* 15.10, p. 2194.

Hattar, S (2003). "Lucas RJ, Mrosovsky N, Thompson S, Douglas RH, Hankins MW, Lem J, Biel M, Hofmann F, Foster RG, and Yau KW". In: *Melanopsin and rod-cone photoreceptive systems account for all major accessory visual functions in mice. Nature* 424, pp. 75–81.

Haverkamp, Silke and Heinz Wässle (2004). "Characterization of an amacrine cell type of the mammalian retina immunoreactive for vesicular glutamate transporter 3". In: *Journal of comparative neurology* 468.2, pp. 251–263.

Helmstaedter, Moritz (2013). "Cellular-resolution connectomics: challenges of dense neural circuit reconstruction". In: *Nature methods* 10.6, pp. 501–507.

Hess, Douglas T, Akio Matsumoto, Sung-Oog Kim, Harvey E Marshall, and Jonathan S Stamler (2005). "Protein S-nitrosylation: purview and parameters". In: *Nature reviews Molecular cell biology* 6.2, pp. 150–166.

Hicks, Stephanie C, F William Townes, Mingxiang Teng, and Rafael A Irizarry (2018). "Missing data and technical variability in single-cell RNA-sequencing experiments". In: *Biostatistics* 19.4, pp. 562–578.

Hinton, Geoffrey E and Sam Roweis (2002). "Stochastic neighbor embedding". In: *Advances in neural information processing systems* 15.

Hirasawa, Hajime, Massimo Contini, and Elio Raviola (2015). "Extrasynaptic release of GABA and dopamine by retinal dopaminergic neurons". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 370.1672, p. 20140186.

Ho, Tin Kam (1995). "Random decision forests". In: *Proceedings of 3rd international conference on document analysis and recognition*. Vol. 1. IEEE, pp. 278–282.

Ho, Tracy, Andrew I Jobling, Ursula Greferath, Trinette Chuang, Archana Ramesh, Erica L Fletcher, and Kirstan A Vessey (2015). "Vesicular expression and release of ATP from dopaminergic neurons of the mouse retina and midbrain". In: *Frontiers in cellular neuroscience* 9, p. 389.

Höfling, Larissa, Klaudia P Szatko, Christian Behrens, Yongrong Qiu, David Alexander Klindt, Zachary Jessen, Gregory S Schwartz, Matthias Bethge, Philipp Berens, Katrin Franke, et al. (2022). "A chromatic feature detector in the retina signals visual context changes". In: *bioRxiv*, pp. 2022–11.

Hubel, David H and Torsten N Wiesel (1959). "Receptive fields of single neurones in the cat's striate cortex". In: *The Journal of physiology* 148.3, p. 574.

Hubel, David H and Torsten N Wiesel (1962). "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex". In: *The Journal of physiology* 160.1, p. 106.

Jacoby, Jason, Amurta Nath, Zachary F Jessen, and Gregory W Schwartz (2018). "A self-regulating gap junction network of amacrine cells controls nitric oxide release in the retina". In: *Neuron* 100.5, pp. 1149–1162.

Kamermans, M and H Spekreijse (1999). "The feedback pathway from horizontal cells to cones: A mini review with a look ahead". In: *Vision research* 39.15, pp. 2449–2468.

Kerr, Rex, Varda Lev-Ram, Geoff Baird, Pierre Vincent, Roger Y Tsien, and William R Schafer (2000). "Optical imaging of calcium transients in neurons and pharyngeal muscle of C. elegans". In: *Neuron* 26.3, pp. 583–594.

Kim, In-Beom, Eun-Jin Lee, Keun-Young Kim, Won-Kyu Ju, Su-Ja Oh, Choun-Ki Joo, and Myung-Hoon Chun (1999). "Immunocytochemical localization of nitric oxide synthase in the mammalian retina". In: *Neuroscience letters* 267.3, pp. 193–196.

Knöpfel, Thomas and Chenchen Song (2019). "Optical voltage imaging in neurons: moving from technology development to practical tool". In: *Nature Reviews Neuroscience* 20.12, pp. 719–727.

Knott, Andrew B and Ella Bossy-Wetzel (2009). "Nitric oxide in health and disease of the nervous system". In: *Antioxidants & redox signaling* 11.3, pp. 541–553.

Knowles, Richard G and Salvador Moncada (1992). "Nitric oxide as a signal in blood vessels". In: *Trends in biochemical sciences* 17.10, pp. 399–402.

Kölsch, Yvonne, Joshua Hahn, Anna Sappington, Manuel Stemmer, António M Fernandes, Thomas O Helmbrecht, Shriya Lele, Salwan Butrus, Eva Laurell, Irene Arnold-Ammer,

et al. (2021). “Molecular classification of zebrafish retinal ganglion cells links genes to cell types to behavior”. In: *Neuron* 109.4, pp. 645–662.

Lettvin, Jerome Y, Humberto R Maturana, Warren S McCulloch, and Walter H Pitts (1959). “What the frog’s eye tells the frog’s brain”. In: *Proceedings of the IRE* 47.11, pp. 1940–1951.

Levick, William R (1967). “Receptive fields and trigger features of ganglion cells in the visual streak of the rabbit’s retina”. In: *The Journal of physiology* 188.3, p. 285.

Liu, Zhuohe, Xiaoyu Lu, Vincent Villette, Yueyang Gou, Kevin L Colbert, Shujuan Lai, Sihui Guan, Michelle A Land, Jihwan Lee, Tensae Assefa, et al. (2022). “Sustained deep-tissue voltage recording using a fast indicator evolved for two-photon microscopy”. In: *Cell* 185.18, pp. 3408–3425.

Marc, Robert E, James R Anderson, Bryan W Jones, Crystal L Sigulinsky, and James S Lauritzen (2014). “The AII amacrine cell connectome: a dense network hub”. In: *Frontiers in neural circuits* 8, p. 104.

Marder, Eve (2012). “Neuromodulation of neuronal circuits: back to the future”. In: *Neuron* 76.1, pp. 1–11.

Marvin, Jonathan S, Bart G Borghuis, Lin Tian, Joseph Cichon, Mark T Harnett, Jasper Akerboom, Andrew Gordus, Sabine L Renninger, Tsai-Wen Chen, Cornelia I Bargmann, et al. (2013). “An optimized fluorescent probe for visualizing glutamate neurotransmission”. In: *Nature methods* 10.2, pp. 162–170.

Masland, Richard H (2001). “The fundamental plan of the retina”. In: *Nature neuroscience* 4.9, pp. 877–886.

Masland, Richard H (2005). “The many roles of starburst amacrine cells”. In: *Trends in Neurosciences* 28.8, pp. 395–396.

Masland, Richard H (2012a). “The neuronal organization of the retina”. In: *Neuron* 76.2, pp. 266–280.

Masland, Richard H (2012b). “The tasks of amacrine cells”. In: *Visual neuroscience* 29.1, pp. 3–9.

Meder, David, Damian Marc Herz, James Benedict Rowe, Stéphane Lehéricy, and Hartwig Roman Siebner (2019). “The role of dopamine in the brain-lessons learned from Parkinson’s disease”. In: *Neuroimage* 190, pp. 79–93.

Mills, Stephen L and Stephen C Massey (1995). “Differential properties of two gap junctional pathways made by AII amacrine cells”. In: *Nature* 377.6551, pp. 734–737.

Nadal-Nicolás, Francisco M, Vincent P Kunze, John M Ball, Brian T Peng, Akshay Krishnan, Gaohui Zhou, Lijin Dong, and Wei Li (2020). “True S-cones are concentrated in the ventral mouse retina and wired for color detection in the upper visual field”. In: *Elife* 9, e56840.

Nawy, Scott and Craig E Jahr (1990). “Suppression by glutamate of cGMP-activated conductance in retinal bipolar cells”. In: *Nature* 346.6281, pp. 269–271.

Neath, Andrew A and Joseph E Cavanaugh (2012). “The Bayesian information criterion: background, derivation, and applications”. In: *Wiley Interdisciplinary Reviews: Computational Statistics* 4.2, pp. 199–203.

Nikonov, Sergei S, Roman Kholodenko, Janis Lem, and Edward N Pugh Jr (2006). "Physiological features of the S-and M-cone photoreceptors of wild-type mice from single-cell recordings". In: *The Journal of general physiology* 127.4, pp. 359–374.

Nir, Izhak, Rashidul Haque, and P Michael Iuvone (2000). "Diurnal metabolism of dopamine in the mouse retina". In: *Brain research* 870.1-2, pp. 118–125.

Nirenberg, Sheila and Markus Meister (1997). "The light response of retinal ganglion cells is truncated by a displaced amacrine circuit". In: *Neuron* 18.4, pp. 637–650.

Nishizaki, T, K Nagai, T Nomura, H Tada, T Kanno, H Tozaki, XX Li, T Kondoh, N Kodama, E Takahashi, et al. (2002). "A new neuromodulatory pathway with a glial contribution mediated via A2a adenosine receptors". In: *Glia* 39.2, pp. 133–147.

Olshausen, Bruno A and David J Field (1996). "Natural image statistics and efficient coding". In: *Network: computation in neural systems* 7.2, pp. 333–339.

Pérez-Fernández, Víctor, Nina Milosavljevic, Annette E Allen, Kirstan A Vessey, Andrew I Jobling, Erica L Fletcher, Paul P Breen, John W Morley, and Morven A Cameron (2019). "Rod photoreceptor activation alone defines the release of dopamine in the retina". In: *Current Biology* 29.5, pp. 763–774.

Prigge, Cameron L, Po-Ting Yeh, Nan-Fu Liou, Chi-Chan Lee, Shih-Feng You, Lei-Lei Liu, David S McNeill, Kylie S Chew, Samer Hattar, Shih-Kuo Chen, et al. (2016). "M1 ipRGCs influence visual function through retrograde signaling in the retina". In: *Journal of Neuroscience* 36.27, pp. 7184–7197.

Qiao, Sheng-Nan, Zhijing Zhang, Christophe P Ribelayga, Yong-Mei Zhong, and Dao-Qi Zhang (2016). "Multiple cone pathways are involved in photic regulation of retinal dopamine". In: *Scientific reports* 6.1, p. 28916.

Qiu, Yongrong, Zhijian Zhao, David Klindt, Magdalena Kautzky, Klaudia P Szatko, Frank Schaeffel, Katharina Rifai, Katrin Franke, Laura Busse, and Thomas Euler (2021). "Natural environment statistics in the upper and lower visual field are reflected in mouse retinal specializations". In: *Current Biology* 31.15, pp. 3233–3247.

Reynolds, Douglas A et al. (2009). "Gaussian mixture models." In: *Encyclopedia of biometrics* 741.659-663.

Rogerson, Luke E, Zhijian Zhao, Katrin Franke, Thomas Euler, and Philipp Berens (2019). "Bayesian hypothesis testing and experimental design for two-photon imaging data". In: *PLoS computational biology* 15.8, e1007205.

Roska, Botond and Frank Werblin (2001). "Vertical interactions across ten parallel, stacked representations in the mammalian retina". In: *Nature* 410.6828, pp. 583–587.

Rosón, Miroslav Román, Yannik Bauer, Ann H Kotkat, Philipp Berens, Thomas Euler, and Laura Busse (2019). "Mouse dLGN receives functional input from a diverse population of retinal ganglion cells with limited convergence". In: *Neuron* 102.2, pp. 462–476.

Roy, Suva and Greg D Field (2019). "Dopaminergic modulation of retinal processing from starlight to sunlight". In: *Journal of pharmacological sciences* 140.1, pp. 86–93.

Sanes, Joshua R and Richard H Masland (2015). "The types of retinal ganglion cells: current status and implications for neuronal classification". In: *Annual review of neuroscience* 38, pp. 221–246.

Santos-Carvalho, Ana, Ana Rita Alvaro, Joao Martins, Antonio Francisco Ambrosio, and Claudia Cavadas (2014). "Emerging novel roles of neuropeptide Y in the retina: from neuromodulation to neuroprotection". In: *Progress in neurobiology* 112, pp. 70–79.

Santos-Carvalho, Ana, António Francisco Ambrósio, and Cláudia Cavadas (2015). "Neuropeptide Y system in the retina: From localization to function". In: *Progress in retinal and eye research* 47, pp. 19–37.

Schröder, Cornelius, David Klindt, Sarah Strauss, Katrin Franke, Matthias Bethge, Thomas Euler, and Philipp Berens (2020a). "System identification with biophysical constraints: A circuit model of the inner retina". In: *Advances in Neural Information Processing Systems* 33, pp. 15439–15450.

Schröder, Sylvia, Nicholas A Steinmetz, Michael Krumin, Marius Pachitariu, Matteo Rizzi, Leon Lagnado, Kenneth D Harris, and Matteo Carandini (2020b). "Arousal modulates retinal output". In: *Neuron* 107.3, pp. 487–495.

Seabrook, Tania A, Timothy J Burbridge, Michael C Crair, and Andrew D Huberman (2017). "Architecture, function, and assembly of the mouse visual system". In: *Annual review of neuroscience* 40, pp. 499–538.

Seung, H Sebastian and Uygar Sümbül (2014). "Neuronal cell types and connectivity: lessons from the retina". In: *Neuron* 83.6, pp. 1262–1272.

Seung, Sebastian (2012). *Connectome: How the brain's wiring makes us who we are*. HMH.

Shabani, H, Mahdi Sadeghi, E Zrenner, DL Rathbun, and Z Hosseinzadeh (2021). "Classification of pseudocalcium visual responses from mouse retinal ganglion cells". In: *Visual Neuroscience* 38, E016.

Shah, Nishal P, Nora Brackbill, Ryan Samarakoon, Colleen Rhoades, Alexandra Kling, Alexander Sher, Alan Litke, Yoram Singer, Jonathon Shlens, and EJ Chichilnisky (2022). "Individual variability of neural computations in the primate retina". In: *Neuron* 110.4, pp. 698–708.

Shekhar, Karthik, Sylvain W Lapan, Irene E Whitney, Nicholas M Tran, Evan Z Macosko, Monika Kowalczyk, Xian Adiconis, Joshua Z Levin, James Nemesh, Melissa Goldman, et al. (2016). "Comprehensive classification of retinal bipolar neurons by single-cell transcriptomics". In: *Cell* 166.5, pp. 1308–1323.

Shi, Qing, Pranjal Gupta, Alexandra K Boukhvalova, Joshua H Singer, and Daniel A Butts (2019). "Functional characterization of retinal ganglion cells using tailored nonlinear modeling". In: *Scientific reports* 9.1, p. 8713.

Smith, Robert G, Michael A Freed, and Peter Sterling (1986). "Microcircuitry of the dark-adapted cat retina: functional architecture of the rod-cone network". In: *Journal of Neuroscience* 6.12, pp. 3505–3517.

Snellman, Josefin and Scott Nawy (2004). "cGMP-dependent kinase regulates response sensitivity of the mouse on bipolar cell". In: *Journal of Neuroscience* 24.29, pp. 6621–6628.

Snyder, Solomon H and David S Bredt (1992). "Biological roles of nitric oxide". In: *Scientific American* 266.5, pp. 68–77.

Ströh, Sebastian, Christian Puller, Sebastian Swirski, Maj-Britt Hölzel, Lea IS van der Linde, Jasmin Segelken, Konrad Schultz, Christoph Block, Hannah Monyer, Klaus Willecke, et al.

(2018). “Eliminating glutamatergic input onto horizontal cells changes the dynamic range and receptive field organization of mouse retinal ganglion cells”. In: *Journal of Neuroscience* 38.8, pp. 2015–2028.

Szatko, Klaudia P, Maria M Korympidou, Yanli Ran, Philipp Berens, Deniz Dalkara, Timm Schubert, Thomas Euler, and Katrin Franke (2020). “Neural circuits in the mouse retina support color vision in the upper visual field”. In: *Nature communications* 11.1, p. 3481.

Taylor, WR and RG Smith (2012). “The role of starburst amacrine cells in visual signal processing”. In: *Visual neuroscience* 29.1, pp. 73–81.

Thoreson, Wallace B and Stuart C Mangel (2012). “Lateral interactions in the outer retina”. In: *Progress in retinal and eye research* 31.5, pp. 407–441.

Thunemann, Martin, Lai Wen, Matthias Hillenbrand, Angelos Vachaviolos, Susanne Feil, Thomas Ott, Xiaoxing Han, Dai Fukumura, Rakesh K Jain, Michael Russwurm, et al. (2013). “Transgenic mice for cGMP imaging”. In: *Circulation research* 113.4, pp. 365–371.

Tran, Nicholas M, Karthik Shekhar, Irene E Whitney, Anne Jacobi, Inbal Benhar, Guosong Hong, Wenjun Yan, Xian Adiconis, McKinzie E Arnold, Jung Min Lee, et al. (2019). “Single-cell profiles of retinal ganglion cells differing in resilience to injury reveal neuroprotective genes”. In: *Neuron* 104.6, pp. 1039–1055.

Trapani, Francesco, Giulia Spampinato, Pierre Yger, and Olivier Marre (2022). “Differences in non-linearities determine retinal cell types”. In: *bioRxiv*, pp. 2022–05.

Tsumamoto, Yuichi, Keisuke Yamashita, Masaya Takumida, Koji Okada, Satoshi Mukai, Makoto Shinya, Hidetoshi Yamashita, and Hiromu K Mishima (2002). “In situ localization of nitric oxide synthase and direct evidence of NO production in rat retinal ganglion cells”. In: *Brain research* 933.2, pp. 118–129.

Van Hateren, J Hans and Arjen van der Schaaf (1998). “Independent component filters of natural images compared with simple cells in primary visual cortex”. In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 265.1394, pp. 359–366.

Van Wyk, Michiel, W Rowland Taylor, and David I Vaney (2006). “Local edge detectors: a substrate for fine spatial vision at low temporal frequencies in rabbit retina”. In: *Journal of Neuroscience* 26.51, pp. 13250–13263.

Vielma, Alex H, Adolfo Agurto, Joaquín Valdés, Adrián G Palacios, and Oliver Schmachtenberg (2014). “Nitric oxide modulates the temporal properties of the glutamate response in type 4 OFF bipolar cells”. In: *PLoS One* 9.12, e114330.

Vielma, Alex H, Mauricio A Retamal, and Oliver Schmachtenberg (2012). “Nitric oxide signaling in the retina: what have we learned in two decades?” In: *Brain research* 1430, pp. 112–125.

Vlasits, Anna L, Thomas Euler, and Katrin Franke (2019). “Function first: classifying cell types and circuits of the retina”. In: *Current opinion in neurobiology* 56, pp. 8–15.

Vogelstein, Joshua T, Adam M Packer, Timothy A Machado, Tanya Sippy, Baktash Babadi, Rafael Yuste, and Liam Paninski (2010). “Fast nonnegative deconvolution for spike train inference from population calcium imaging”. In: *Journal of neurophysiology* 104.6, pp. 3691–3704.

Völgyi, Béla, Samir Chheda, and Stewart A Bloomfield (2009). "Tracer coupling patterns of the ganglion cell subtypes in the mouse retina". In: *Journal of Comparative Neurology* 512.5, pp. 664–687.

Volkow, Nora D and Marisela Morales (2015). "The brain on drugs: from reward to addiction". In: *Cell* 162.4, pp. 712–725.

Wang, Guo-Yong, Lauren C Liets, and Leo M Chalupa (2003). "Nitric oxide differentially modulates ON and OFF responses of retinal ganglion cells". In: *Journal of neurophysiology* 90.2, pp. 1304–1313.

Warwick, Rebekah A, Alina S Heukamp, Serena Riccitelli, and Michal Rivlin-Etzion (2023). "Dopamine differentially affects retinal circuits to shape the retinal code". In: *The Journal of physiology* 601.7, pp. 1265–1286.

Warwick, Rebekah A, Serena Riccitelli, Alina S Heukamp, Hadar Yaakov, Lea Ankri, Jonathan Mayzel, Noa Gilead, Reut Parness-Yossifon, and Michal Rivlin-Etzion (2022). "Top-down modulation of the retinal code via histaminergic neurons of the hypothalamus". In: *bioRxiv*, pp. 2022–04.

Wässle, Heinz (2004). "Parallel processing in the mammalian retina". In: *Nature Reviews Neuroscience* 5.10, pp. 747–757.

Wässle, Heinz, Christian Puller, Frank Müller, and Silke Haverkamp (2009). "Cone contacts, mosaics, and territories of bipolar cells in the mouse retina". In: *Journal of Neuroscience* 29.1, pp. 106–117.

Weng, Shijun, Wenzhi Sun, and Shigang He (2005). "Identification of ON–OFF direction-selective ganglion cells in the mouse retina". In: *The Journal of physiology* 562.3, pp. 915–923.

Werblin, Frank S and John E Dowling (1969). "Organization of the retina of the mudpuppy, Necturus maculosus. II. Intracellular recording." In: *Journal of neurophysiology* 32.3, pp. 339–355.

Wienbar, Sophia and Gregory William Schwartz (2022). "Differences in spike generation instead of synaptic inputs determine the feature selectivity of two retinal cell types". In: *Neuron* 110.13, pp. 2110–2123.

Witkovsky, Paul (2004). "Dopamine and retinal function". In: *Documenta ophthalmologica* 108, pp. 17–39.

Witkovsky, Paul and Allen Dearry (1991). "Functional roles of dopamine in the vertebrate retina". In: *Progress in retinal research* 11, pp. 247–292.

Wu, Samuel M, Fan Gao, and Bruce R Maple (2000). "Functional architecture of synapses in the inner retina: segregation of visual signals by stratification of bipolar cell axon terminals". In: *Journal of Neuroscience* 20.12, pp. 4462–4470.

Wässle, Heinz and Brian B Boycott (1991). "Functional architecture of the mammalian retina". In: *Physiological reviews* 71.2, pp. 447–480.

Yan, Wenjun, Mallory A Laboulaye, Nicholas M Tran, Irene E Whitney, Inbal Benhar, and Joshua R Sanes (2020). "Mouse retinal cell atlas: molecular identification of over sixty amacrine cell types". In: *Journal of Neuroscience* 40.27, pp. 5177–5195.

Yang, Albert C and Shih-Jen Tsai (2017). "New targets for schizophrenia treatment beyond the dopamine hypothesis". In: *International journal of molecular sciences* 18.8, p. 1689.

Yau, King-Wai and Roger C Hardie (2009). "Phototransduction motifs and variations". In: *Cell* 139.2, pp. 246–264.

Yazulla, Stephen (2008). "Endocannabinoids in the retina: from marijuana to neuroprotection". In: *Progress in retinal and eye research* 27.5, pp. 501–526.

Zhang, Dao-Qi, Kwoon Y Wong, Patricia J Sollars, David M Berson, Gary E Pickard, and Douglas G McMahon (2008). "Intraretinal signaling by ganglion cell photoreceptors to dopaminergic amacrine neurons". In: *Proceedings of the National Academy of Sciences* 105.37, pp. 14181–14186.

Zhao, Xiwu, Kwoon Y Wong, and Dao-Qi Zhang (2017). "Mapping physiological inputs from multiple photoreceptor systems to dopaminergic amacrine cells in the mouse retina". In: *Scientific reports* 7.1, p. 7920.

Zhao, Zhijian, David A Klindt, André Maia Chagas, Klaudia P Szatko, Luke Rogerson, Dario A Protti, Christian Behrens, Deniz Dalkara, Timm Schubert, Matthias Bethge, et al. (2020). "The temporal structure of the inner retina at a single glance". In: *Scientific reports* 10.1, p. 4399.

Zou, Hui, Trevor Hastie, and Robert Tibshirani (2006). "Sparse principal component analysis". In: *Journal of computational and graphical statistics* 15.2, pp. 265–286.

Publications

1. Yongrong Qiu, David A. Klindt, Klaudia P. Szatko, Dominic Gonschorek, Larissa Hoefling, Timm Schubert, Laura Busse, Matthias Bethge, Thomas Euler (2023). *Efficient coding of natural scenes improves neural system identification*. PLOS Computational Biology, 19(4), e1011037. **published**.

2. Dominic Gonschorek*, Larissa Hoefling*, Klaudia P. Szatko, Katrin Franke, Timm Schubert, Benjamin Dunn, Philipp Berens, David A. Klindt, Thomas Euler (2021). *Removing inter-experimental variability from functional data in systems neuroscience*. Advances in Neural Information Processing Systems (NeurIPS 2021), 34, 3706-3719. **published**.

3. Dominic Gonschorek, Tom Schwerd-Kleine, Jonathan Oesterle, Ryan Arlinghaus, Zhijan Zhao, Timm Schubert, Thomas Euler (2023). *Nitric oxide mediates different effects in mouse retinal ganglion cells*. **in preparation**.

Efficient coding of natural scenes improves neural system identification

Yongrong Qiu, David A. Klindt, Klaudia P. Szatko, Dominic Gonschorek, Larissa Hoefling, Timm Schubert, Laura Busse, Matthias Bethge, Thomas Euler (2023). *PLOS Computational Biology, 19(4), e1011037*. **published**, see Appendix, Study I.

Abstract: Neural system identification aims at learning the response function of neurons to arbitrary stimuli using experimentally recorded data, but typically does not leverage normative principles such as efficient coding of natural environments. Visual systems, however, have evolved to efficiently process input from the natural environment. Here, we present a normative network regularization for system identification models by incorporating, as a regularizer, the efficient coding hypothesis, which states that neural response properties of sensory representations are strongly shaped by the need to preserve most of the stimulus information with limited resources. Using this approach, we explored if a system identification model can be improved by sharing its convolutional filters with those of an autoencoder which aims to efficiently encode natural stimuli. To this end, we built a hybrid model to predict the responses of retinal neurons to noise stimuli. This approach did not only yield a higher performance than the "stand-alone" system identification model, it also produced more biologically plausible filters, meaning that they more closely resembled neural representation in early visual systems. We found these results applied to retinal responses to different artificial stimuli and across model architectures. Moreover, our normatively regularized model performed particularly well in predicting responses of direction-of-motion sensitive retinal neurons. The benefit of natural scene statistics became marginal, however, for predicting the responses to natural movies. In summary, our results indicate that efficiently encoding environmental inputs can improve system identification models, at least for noise stimuli, and point to the benefit of probing the visual system with naturalistic stimuli.

Contributions: Methodology and software development of the retinal ganglion cell type classifier, data visualization, writing and editing of the manuscript.

Removing Inter-Experimental Variability from Functional Data in Systems Neuroscience

Dominic Gonschorek*, Larissa Hoefling*, Klaudia P. Szatko, Katrin Franke, Timm Schubert, Benjamin Dunn, Philipp Berens, David Klindt, Thomas Euler (2021). *Advances in Neural Information Processing Systems (NeurIPS 2021), 34, 3706-3719*. **published**, see Appendix, Study II.

Abstract: Integrating data from multiple experiments is common practice in systems neuroscience but it requires inter-experimental variability to be negligible compared to the biological signal of interest. This requirement is rarely fulfilled; systematic changes between experiments can drastically affect the outcome of complex analysis pipelines. Modern machine learning approaches designed to adapt models across multiple data domains offer flexible ways of removing inter-experimental variability where classical statistical methods often fail. While applications of these methods have been mostly limited to single-cell genomics, in this work, we develop a theoretical framework for domain adaptation in systems neuroscience. We implement this in an adversarial optimization scheme that removes inter-experimental variability while preserving the biological signal. We compare our method to previous approaches on a large-scale dataset of two-photon imaging recordings of retinal bipolar cell responses to visual stimuli. This dataset provides a unique benchmark as it contains biological signal from well-defined cell types that is obscured by large inter-experimental variability. In a supervised setting, we compare the generalization performance of cell type classifiers across experiments, which we validate with anatomical cell type distributions from electron microscopy data. In an unsupervised setting, we remove inter-experimental variability from data which can then be fed into arbitrary downstream analyses. In both settings, we find that our method achieves the best trade-off between removing inter-experimental variability and preserving biological signal. Thus, we offer a flexible approach to remove inter-experimental variability and integrate datasets across experiments in systems neuroscience. Code available at `https://github.com/eulerlab/rave`.

Contributions: Conceptualization, data analysis, co-development of the theoretical framework, data visualization, writing and editing of the manuscript, manuscript revision.

Nitric oxide mediates different effects in mouse retinal ganglion cells

Dominic Gonschorek, Tom Schwerd-Kleine, Jonathan Oesterle, Ryan Arlinghaus, Zhijan Zhao, Timm Schubert, Thomas Euler (2023). **in preparation**, see Appendix, Study III.

Abstract: Neural system identification aims at learning the response function of neurons to arbitrary stimuli using experimentally recorded data, but typically does not leverage normative principles such as efficient coding of natural environments. Visual systems, however, have evolved to efficiently process input from the natural environment. Here, we present a normative network regularization for system identification models by incorporating, as a regularizer, the efficient coding hypothesis, which states that neural response properties of sensory representations are strongly shaped by the need to preserve most of the stimulus information with limited resources. Using this approach, we explored if a system identification model can be improved by sharing its convolutional filters with those of an autoencoder which aims to efficiently encode natural stimuli. To this end, we built a hybrid model to predict the responses of retinal neurons to noise stimuli. This approach did not only yield a higher performance than the "stand-alone" system identification model, it also produced more biologically plausible filters, meaning that they more closely resembled neural representation in early visual systems. We found these results applied to retinal responses to different artificial stimuli and across model architectures. Moreover, our normatively regularized model performed particularly well in predicting responses of direction-of-motion sensitive retinal neurons. The benefit of natural scene statistics became marginal, however, for predicting the responses to natural movies. In summary, our results indicate that efficiently encoding environmental inputs can improve system identification models, at least for noise stimuli, and point to the benefit of probing the visual system with naturalistic stimuli.

Contributions: Performed imaging experiments, preprocessing of data, data analysis pipeline, data analysis, data visualization, writing and editing of the manuscript.

Acknowledgments

I would like to thank everyone with deep gratitude who helped, guided, mentored and has lent me an ear throughout this exciting, often challenging, but primarily interesting time on my journey to become a Ph.D. I thank Thomas Euler for being a great supervisor and mentor, who preceded as a good example of mentorship. He always supported me at all stages of my Ph.D. and also let me grow and develop as a person and young researcher. I thank Jan Benda and Clint Makino for their great scientific input and their time on my thesis advisory committee. I thank Timm Schubert for his great and on-point scientific inputs, feedback and many walks and conversations in the woods around Tübingen. I thank David A. Klindt for his collaboration on my first publication, his excellent supervision and his belief in me. I thank Jonathan Oesterle for his support towards the end of my Ph.D. and the fun hikes and trips. I thank Klaudia P. Szatko for teaching me the experiments. I thank Zhijian Zhao for this excellent technical support on the setups and his scientific input. I thank Georgios Kafetzis for the many fruitful conversations and the early morning walks through Tübingen. I thank Merle Harrer for her amazing technical assistance in the lab. I thank Gordon Eske for his technical assistance and great conversations during the beginning of my Ph.D. I thank Valeska Botzenhardt for her administrative support. I thank Marili Korympidou, Larissa Höfling, Nadine Dyszkant, Tom Schwerd-Kleine, and the rest of the current and former EulerLab members for their fruitful discussions, countless coffee talks and much more. I thank Philipp Berens and his lab for great discussions and support on computational questions. I thank the GRK and all its members for providing an intellectual and supportive environment throughout my Ph.D. Additionally, I'd like to thank Robert Feil and Malte Roeßling for providing their mouse lines and scientific input. I thank Olivier Marre and his entire lab for hosting me in Paris and allowing me to learn new techniques and having a blast time. A very special and huge thanks to Andrés Covelli, Philipp Hentze and Christian Rieger (and his family) for always being there, supporting me during good and dark times, for all the fun and honest conversations. Without you, I wouldn't be the person I am right now. I thank Stefan Brauer for his long friendship, early mentorship and the countless conversations that sparked my interest in neuroscience in the first place and kicked-off my journey. I thank Wil Brown for his long friendship, all the fun moments and trips. You are an amazing coach, human being and friend who had a huge contribution in my personal development. In this context, I also thank our mentorship group and the lovely people from the London Movement Group for always giving me a good time and laugh. I also thank Lisa Schmors for her scientific and emotional support, even when times have not been easy. I thank Meike Rogalla for the many uplifting and honest conversations during the time of my Ph.D. I thank Fabienne Windel for being a good listener and external source of unbiased input as a Ph.D. student. I thank all my other past and current friends and companions that lead me along my life and this way. Finally, I thank my family for their unconditional support and love during all phases, which has brought me to exactly this point in my life.

Appendix

Study I: Efficient coding of natural scenes improves neural system identification:

Published as "Qiu, Yongrong, et al. 'Efficient coding of natural scenes improves neural system identification'; PLOS Computational Biology 19.4 (2023): e1011037."

PLOS COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Efficient coding of natural scenes improves neural system identification

Yongrong Qiu[1,2,3], David A. Klindt[1,2,4], Klaudia P. Szatko[1,2,3,5], Dominic Gonschorek[1,2,6], Larissa Hoefling[1,2,5], Timm Schubert[1,2], Laura Busse[7,8], Matthias Bethge[2,5,9], Thomas Euler[1,2,5]*

1 Institute for Ophthalmic Research, U Tübingen, Tübingen, Germany, **2** Centre for Integrative Neuroscience (CIN), U Tübingen, Tübingen, Germany, **3** Graduate Training Centre of Neuroscience (GTC), International Max Planck Research School, U Tübingen, Tübingen, Germany, **4** Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway, **5** Bernstein Center for Computational Neuroscience, Tübingen, Germany, **6** Research Training Group 2381, U Tübingen, Tübingen, Germany, **7** Division of Neurobiology, Faculty of Biology, LMU Munich, Planegg-Martinsried, Germany, **8** Bernstein Center for Computational Neuroscience, Planegg-Martinsried, Germany, **9** Institute for Theoretical Physics, U Tübingen, Tübingen, Germany

* thomas.euler@cin.uni-tuebingen.de

Citation: Qiu Y, Klindt DA, Szatko KP, Gonschorek D, Hoefling L, Schubert T, et al. (2023) Efficient coding of natural scenes improves neural system identification. PLoS Comput Biol 19(4): e1011037. https://doi.org/10.1371/journal.pcbi.1011037

Editor: Roland W. Fleming, University of Giessen, GERMANY

Received: December 16, 2022

Accepted: March 20, 2023

Published: April 24, 2023

Data Availability Statement: The data and code for this work are available at https://doi.org/10.5281/zenodo.7656868 and https://github.com/eulerlab/hybrid-model, respectively.

Funding: This work was supported by the German Research Foundation (DFG; SFB 1233, Robust Vision: Inference Principles and Neural Mechanisms, projects 10 and 12, project number 276693517 to L.B., M.B., and T.E.; GRK2381, project number 335549539 to T.E.), the Germany's Excellence Strategy (EXC 2064, project number 390727645 to M.B.), and the European Union's

Abstract

Neural system identification aims at learning the response function of neurons to arbitrary stimuli using experimentally recorded data, but typically does not leverage normative principles such as efficient coding of natural environments. Visual systems, however, have evolved to efficiently process input from the natural environment. Here, we present a normative network regularization for system identification models by incorporating, as a regularizer, the *efficient coding* hypothesis, which states that neural response properties of sensory representations are strongly shaped by the need to preserve most of the stimulus information with limited resources. Using this approach, we explored if a system identification model can be improved by sharing its convolutional filters with those of an autoencoder which aims to efficiently encode natural stimuli. To this end, we built a hybrid model to predict the responses of retinal neurons to noise stimuli. This approach did not only yield a higher performance than the "stand-alone" system identification model, it also produced more biologically plausible filters, meaning that they more closely resembled neural representation in early visual systems. We found these results applied to retinal responses to different artificial stimuli and across model architectures. Moreover, our normatively regularized model performed particularly well in predicting responses of direction-of-motion sensitive retinal neurons. The benefit of natural scene statistics became marginal, however, for predicting the responses to natural movies. In summary, our results indicate that efficiently encoding environmental inputs can improve system identification models, at least for noise stimuli, and point to the benefit of probing the visual system with naturalistic stimuli.

Author summary

Computational models use experimental data to learn stimulus-response functions of neurons, but they are rarely informed by normative coding principles, such as the idea

Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant (agreement No 674901, to T.S., M.B., and T.E.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

that sensory neural systems have evolved to efficiently process natural stimuli. We here introduce a novel method to incorporate natural scene statistics to predict responses of retinal neurons to visual stimuli. We show that considering efficient representations of natural scenes improves the model's predictive performance and produces biologically-plausible receptive fields, at least for responses to noise stimuli. Generally, our approach provides a promising framework to test various (normative) coding principles using experimental data for understanding the computations of biological neural networks.

Introduction

In the past years, advances in experimental techniques enabled detailed, large-scale measurements of activity at many levels of sensory processing [1]. As a consequence, *neural system identification* (SI) approaches have flourished (Fig 1a top). They empirically fit the stimulus-response (transfer) function of neurons based on experimentally recorded data [2–4]. A classic example is the generalized linear model (GLM, [2, 5]), which consists of a linear filter as a first order approximation of a neuron's response function (i.e., its receptive field; [6]), followed by a point-wise nonlinear function for the neuron's output. To account for additional non-linearities (e.g., [7, 8]), several extensions, such as linear-nonlinear cascades [9, 10], have been proposed. More recently, deep neural network-based SI approaches inspired by the hierarchical processing along the visual pathway [11, 12] have been developed (reviewed in [13–17]). While SI methods became particularly successful in predicting responses of visual neurons [18–22], they often require large amounts of training data and, more critically, do rarely consider adaptions to the natural environment.

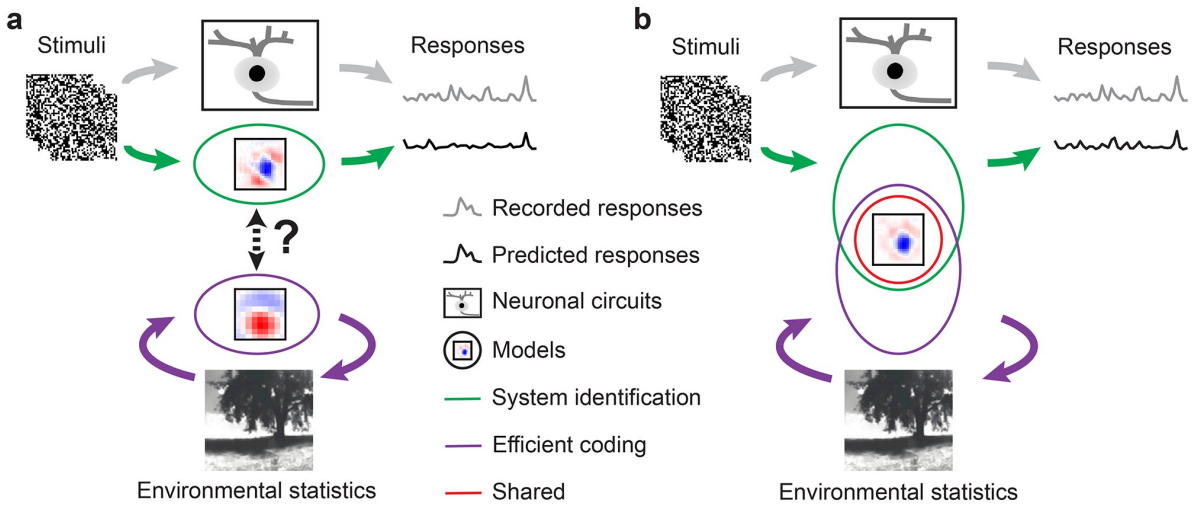

Fig 1. Illustration of our hybrid model combining SI and EC. a. Illustration of two common approaches to studying visual systems: system identification, symbolized by the green-labeled branch, aims at predicting responses of neuronal circuits (black rectangle) to specific stimuli, whereas efficient coding (purple-labeled branch) seeks working out principles of the visual system based on environmental statistics. As these two approaches are rarely combined in a single modeling framework, their potential synergies remain largely unexplored. **b.** Our hybrid modeling approach combines system identification (green) and efficient coding (purple) in a single model with shared filters (red circle) to predict neural responses to arbitrary visual stimuli.

https://doi.org/10.1371/journal.pcbi.1011037.g001

However, like other senses, vision has evolved to promote a species' survival in its natural environment [23], which is thought to have driven visual circuits to efficiently represent information under a number of constraints, including metabolic limits and space restrictions [24, 25]. As a consequence, the visual system has adapted to natural statistics, as shown, for example, by the fact that the distribution of orientation preferences of visual neurons mirrors the dominance of cardinal orientations in natural scenes [26–28].

Such adaptations are at the heart of *efficient coding* (EC) approaches (Fig 1a bottom): They derive algorithmic principles underlying neural systems from the statistical properties of natural stimuli and by incorporating biological constraints [15, 24, 25, 29–31]. Here, one popular strategy starts from the assumption that early visual processing serves to decorrelate the redundant signals in natural environments [32–34]. This theory can reproduce feature selectivity, e.g., difference-of-Gaussian (DoG) kernels that have similar receptive field (RF) properties as retinal ganglion cells (RGCs) [35]. Recently, deep neural network-augmented EC approaches were proposed, such as convolutional autoencoders [36, 37], which are trained to optimally reconstruct inputs in the presence of an information "bottleneck" (i.e., from a constrained latent representation). Such convolutional autoencoders have been shown to yield center-surround spatial RFs with similar properties as those observed in RGCs when encoding either pink ($1/f$) noise or natural scenes [38, 39]. A downside of EC is that it is not always straightforward to experimentally measure coding efficiency and feature selectivity predicted by these approaches in neural systems (discussed in [40, 41]) and, hence, the interpretation of EC models with respect to the biological underpinnings remains challenging.

Notably, the intersection between EC and SI has long remained largely unexplored but lately shifted into focus. For instance, Młynarski and colleagues recently proposed a theoretical framework incorporating normative theories for statistical inference on simulated or pre-fit neural data [42]. Their framework enables conducting rigorous statistical hypothesis tests of coding principles, but has not yet been applied to predicting neural responses to arbitrary stimuli with minimal assumptions.

Here, we tested whether the EC hypothesis can serve as a useful regularization for learning the response functions of neurons. To do so, we built a hybrid model combining a SI branch with an EC branch, forced the two branches to share filters (Fig 1b), and asked if knowledge about natural scene statistics could help predicting retinal responses. To this end, we experimentally recorded Ca^{2+} signals of neurons in the mouse retina while presenting it with noise stimuli. We then used the responses to train the SI branch, which aimed to predict retinal responses. We used natural movies that we recorded in mouse habitats outdoors to train the EC branch, which aimed to represent natural scenes efficiently [39]. We found a synergy between neural prediction and natural scene statistics: First, for noise stimuli, the hybrid approach had a better predictive performance than a pure SI approach. Second, compared to the SI model, the hybrid model produced filters with a clearer center-surround RF structure, akin to RFs at early visual processing stage. However, we did not observe such a synergy for the prediction of responses to natural movies. Our results demonstrate that predicting sensory responses, in particular to noise stimuli, benefits from considering adaptations to the natural environment, and thus highlights the benefits of naturalistic stimuli for vision research.

Materials and methods

Ethics statement

All procedures were performed in accordance with the law on animal protection issued by the German Federal Government (Tierschutzgesetz) and approved by the institutional animal welfare committee of the University of Tübingen.

Animal procedures and retinal activity recordings

Animal procedures. We used n = 5, 5–9 weeks old female C57BL/6 mice (wild-type; JAX 000664, Jackson Laboratory, USA). Due to the exploratory nature of our study, we did not use any statistical methods to predetermine sample size, nor did we perform blinding or randomization. Animals were housed under a standard light-dark (12h:12h) cycle. All procedures were carried out under very dim red illumination (>650 nm). Prior to the start of the experiment, animals were dark-adapted for $\geq$1 h, then anesthetized with isoflurane (Baxter, Germany), and killed by cervical dislocation.

The eyes were enucleated and hemisected in carboxygenated (95% O_2, 5% CO_2) artificial cerebrospinal fluid (ACSF) solution containing (in mM): 125 NaCl, 2.5 KCl, 2 $CaCl_2$, 1 $MgCl_2$, 1.25 NaH_2PO_4, 26 $NaHCO_3$, 20 glucose, and 0.5 l-glutamine (pH 7.4). Next, the retina was flat-mounted onto an Anodisc (#13, 0.1 μm pore size, GE Healthcare, Germany) with the ganglion cell layer (GCL) facing up. To uniformly label the GCL cells, bulk electroporation was performed with the fluorescent Ca^{2+} indicator Oregon-Green BAPTA-1 (OGB-1; Invitrogen, Germany), as described earlier [43, 44], using 4-mm plate electrodes (CUY700P4E/L, Xceltis, Germany) and 9 pulses ($\sim$9.2 V, 100 ms pulse width at 1 Hz). After electroporation, the tissue was immediately moved to the microscope's recording chamber, where it was continuously perfused with carboxygenated ACSF at $\sim$36˚C and left to recover for $\sim$30 min before recordings started. Additionally, Sulforhodamine-101 (SR101, Invitrogen, Germany) was added to the ACSF ($\sim$0.1 μM final concentration) to visualize blood vessels and identify damaged cells.

Two-photon Ca^{2+} recordings and light stimulation. We recorded light stimulus-evoked Ca^{2+} signals in GCL cells of the explanted mouse retina using a MOM-type two-photon (2P) microscope [45, 46] from Sutter Instruments (purchased from Science Products, Germany), as described earlier [44, 47]. In brief, the microscope was powered by a mode-locked Ti: Sapphire laser (MaiTai-HP DeepSee, Newport Spectra-Physics, Germany) at 927 nm. Two detection pathways allowed simultaneously recording of OGB-1 and SR101 fluorescence (HQ 510/84 and HQ 630/60, respectively; both Chroma/AHF, Germany) through a 16x water immersion objective (CFI75 LWD16 /0.8W, DIC N2, Nikon, Germany). A custom-written software (ScanM, by M. Müller and T.E.) running under IGOR Pro 6.3 for Windows (Wavemetrics, USA) was used to acquire time-lapsed (64x64 pixels) image scans at a frame rate of 7.8125 Hz. Higher resolution images were acquired using 512x512 pixel scans. Additionally, to register the scan field positions, the outline of the retina and the optic disc were traced.

The retinas were presented with color noise stimulus using a visual stimulator tuned to the spectral sensitivities of mice [48]. This stimulus consisted of independent binary dense noise (28x28 pixel frames, each pixel covering $(0.83^{\circ})^2$ of visual angle) in the UV and green stimulator channels at 5 or 30 Hz. The stimulus contained 5 different training sequences (96 s each) interspersed with 6 repeats of a 10 s test sequence (S1(a) Fig).

In total, we used four data sets for modeling: (*i*) responses of n = 96 GCL neurons to 5-Hz noise recorded in dorsal retina (n = 2 eyes); (*ii*) responses of n = 427 GCL neurons to 5-Hz noise recorded ventrally (n = 5 eyes); in this dataset, we also presented two other stimuli: a full-field chirp (700 μm in diameter) and a moving bar stimulus (300x1,000 μm bright bar moving at 8 directions at 1 mm/s). The responses to these latter stimuli were used to functionally classify the recorded GCL neurons [47]. (*iii*) n = 64 GCL neurons to 30-Hz noise recorded ventrally (n = 2 eyes). (*iv*) n = 86 GCL neurons to 30-Hz natural movie recorded ventrally (n = 1 eye). All cell numbers are after quality control (see below).

Data preprocessing and analysis. For each cell, we calculated a quality index (*QI*, with $0 \leq QI \leq 1$) for its responses to each stimulus type as follows:

$$QI = \mathrm{Var}[\mathrm{E}[C]_r]_t / \mathrm{E}[\mathrm{Var}[C]_t]_r \quad (1)$$

where C is a t-by-r response matrix (time samples, *t*, by repetitions, *r*). The higher *QI*, the more reliable the response and the higher the signal-to-noise ratio. For the noise stimulus, QI_{noise} was determined based on the test sequence responses. For the following analysis, we only used cells with $QI_{noise} > 0.25$; in case chirp and moving bar responses were also recorded, neurons had to fulfill $QI_{chirp} > 0.35$ or $QI_{bar} > 0.6$ to be included.

In case of the noise stimulus, we preprocessed each cell's Ca^{2+} signal by Z-normalizing the raw traces and matching sampling frequency of the recording (7.8125 Hz) to the stimulus frequency (5 or 30 Hz) via linear interpolation. Then, the traces were detrended using a high-pass filter (> 0.1 Hz) and their 1st order derivatives were calculated, with negative values set to zero. We used the average of a cell's responses to the 6 test sequence repeats as ground truth. Excluding the test sequences, we had per cell a total of 480 s of data, of which we used 440 s (∼91%) for training and the remaining 40 s (∼9%) for validation (i.e., to pick the hyperparameters of the SI model, see below).

For the responses to the natural movie stimulus, we used the average of a cell's responses to the 3 test sequence repeats as ground truth. Excluding the test sequences, we had per cell a total of 540 s of data, of which we used 433 s for training and the remaining 107 s for validation. Note that as input for the models, we down-sampled the natural movie stimulus to 36x32 pixel frames to match it to the resolution of the noise stimulus.

For chirp and moving bar responses, we first detrended the traces and then normalized them to [0, 1] [44]. Using these responses, the cells were classified to different functional groups [47] using RGC type classifier (see below).

To estimate the directional tuning from the moving bar responses, we first performed singular value decomposition (SVD) on the mean response matrix, resulting in a temporal and a directional component. We then summed the directional vectors in 2D planes and used the resulting vector length as direction selectivity index. Next, by shuffling trial labels and computing the tuning curve for 1,000 times (permutation test), we got the null distribution (no directional tuning). The percentile of true vector length was used as p-value of directional tuning [47]. Here, we considered cells with $p < 0.05$ as direction-selective (DS) and the remaining ones as non-DS.

RGC type classifier. To predict the functional type of GCL cells, we used a Random Forest Classifier (RFC; [49]), which was trained on a published mouse dataset [47]. In that study, features were extracted from the responses to different visual stimuli (e.g., chirp and moving bar) and used to cluster GCL cells into 32 RGC types and 14 additional dAC types. Here, we learned a mapping *f* from response features (20 features from responses to chirp, ϕ_{chirp} and 8 features from responses to moving bar stimulus, ϕ_{mb}) and two additional parameters $\Theta = \{\theta_{soma}, \theta_{DS}\}$ to functional cell type labels *L* by training a RFC for the dataset from [47]:

$$f : (\phi_{chirp}, \phi_{bar}, \Theta) \mapsto L \quad (2)$$

where θ_{soma} denotes soma size to distinguish between alpha and non-alpha RGC types and θ_{DS} denotes p-value of permutation test for direction selectivity to distinguish between DS and non-DS RGC types.

We fit the RFC on a subset of data from [47] and validated its performance on a held-out test dataset. The classifier had a prediction accuracy of ∼76% on a held-out test dataset (S5 Fig). To apply the trained classifier to our newly recorded dataset, we projected the RGC

responses (normalized to [−1, 1]) into the feature space described in [47] by computing the dot product between the response and the feature matrices. We used the RFC implementation provided by the python package `scikit-learn` [50] to train the classifier.

2D models

Stand-alone SI model (2D). As baseline model to predict the responses of neurons to the noise stimulus, we employed a stand-alone SI model (supervised learning), in which we used factorized spatial and temporal convolutional filters (cf. Fig 2a; [51, 52]). This SI model consisted of one spatial convolutional layer (16x2x1x9x9, output channels x input channels x depth x image width x image height), one temporal convolutional layer (16x16x8x1x1, with 8 stimulus frames preceding an event for noise; 16x16x50x1x1, with 50 stimulus frames preceding an event for natural movie), and—after flattening the spatial dimension—one fully connected layer (FC; 96x6,400 for noise stimulus, and 86x10,752 for natural movie, output x input channels), followed by an exponential function. No padding was used. We tested different filter channel numbers and found the number = 16, 24, 32 had similar performance (higher than number = 8) on our datasets. Then we picked a relatively small number = 16 as the autoencoder models desired large memory in hidden layers (see below). The loss function was defined as:

$$L_{SI} = \sum_i (\hat{\vec{r}}_i - \vec{r}_i \log \hat{\vec{r}}_i) + \alpha_1 \|\overrightarrow{w_{cs}}\|_2 + \alpha_2 \|\overrightarrow{w_{ct}}\|_2 + \beta \|\overrightarrow{w_f}\|_1 \quad (3)$$

Here, the first term is the Poisson loss between predicted responses ($\hat{\vec{r}}_i$) and ground truth ($\overrightarrow{r_i}$) (with i denoting the neuron index), the second term is the L2 penalty on the weights of the spatial convolutional filters ($\overrightarrow{w_{cs}}$) with hyperparameter α_1, the third term is the L2 penalty on the weights of temporal convolutional filters ($\overrightarrow{w_{ct}}$) with hyperparameter α_2, and the last term is the L1 penalty on the FC layer ($\overrightarrow{w_f}$) with hyperparameter β. We note that, compared to the EC branch of hybrid model, penalty on filter weights could be seen as an implicit form of efficient energy coding, limiting synaptic transmission and generating kernels akin to representations in early visual system ([53]).

After performing a grid search for the three hyperparameters, we picked $\alpha_1 = 10$, $\alpha_2 = 10$, $\beta = 1/16$ which yielded the best performance on the validation data. After training, we estimated the neurons' spatio-temporal RF filters by computing gradients for each neuron, starting with a blank image sequence as input. These gradients represent the first-order approximation of the input that maximizes the neuron's activation [6]. For visualization, we extracted the spatial and temporal RFs via SVD.

As a metric of biological plausibility, we calculated the coefficient of determination (R-squared; [0, 1]) of fitting 2D Gaussian distributions to the spatial (component of) the convolutional filters. We set the R-squared value to 0 if the sigma of the fitted Gaussian was larger than the size of the filter (i.e., 9 pixels). We calculated this fit quality for the filter of the chromatic channel with the dominant response. Because the mouse retina is divided into a more green-sensitive dorsal and a more UV-sensitive ventral retina (e.g., [44]), this meant that for dorsal neurons we only determined the R-squared for filters for the green stimulus channel, and for ventral neurons for the UV stimulus channel.

SI-PCA model (2D). The spatial convolutional filters of the SI-PCA model were composed from PCA basis functions (W). The model was trained to learn the weights of these basis functions. The filters were produced by performing PCA transformation on natural images recorded in mouse habitats [39]:

$$W = U^T \quad (4)$$

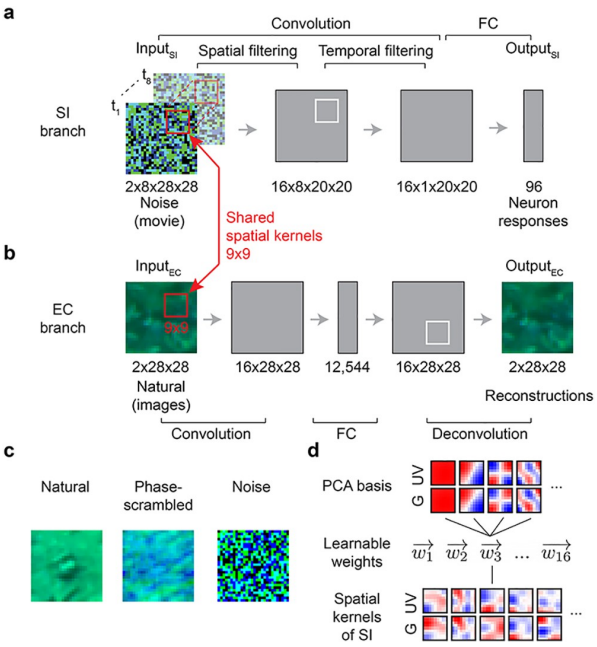

Fig 2. Hybrid model with shared spatial filters. a,b. Schemata of SI model (a) and EC model (b) from [39]. The SI model branch consists of spatial and temporal convolutional layers, a fully connected (FC) layer and a nonlinear layer (see Methods). The EC model branch is a convolutional autoencoder, consisting of an encoder and a decoder network. In the hybrid model, the two branches were trained in parallel with shared spatial filters (all spatial filters were shared; red). $Input_{SI}$: 8-frame UV-green noise ($t_1 \ldots t_8$); $Output_{SI}$: predicted GCL cell Ca^{2+} responses; $Input_{EC}$: UV-green natural images; $Output_{EC}$: reconstructed $Input_{EC}$. **c.** Example for the different inputs (natural images, phase-scrambled natural images, and noise) for the EC branch in hybrid models (*hybrid-natural*, *hybrid-pha-scr*, *hybrid-noise*). **d.** Using PCA filters as basis vectors for spatial convolutional filters of the SI model; *SI-PCA* learned 16 weight vectors ($\overrightarrow{w_1} \ldots \overrightarrow{w_{16}}$) with same vector length as the number of PCA basis elements.

https://doi.org/10.1371/journal.pcbi.1011037.g002

where U contains the eigenvectors of the covariance matrix of the centered data in each column.

For example, when using 4 PCA bases, the shape of learnable weight matrix was 16x4 (channel number x basis number), the shape of PCA bases was 4x2x1x9x9 (basis number x chromatic channel x depth x image width x image height), and the resulted spatial filter had the shape of 16x2x1x9x9. We varied the number of used basis (hyperparameter) and selected the one which achieved the best performance on validation data (S1(b) and S3(b) Figs).

SI-DCT model (2D). For the SI-DCT model, its spatial convolutional filters were composed from DCT basis functions, which were defined as:

$$F(u,v) = \alpha(u)\alpha(v)\cos\left[\frac{(2i+1)\pi}{2N}u\right]\cos\left[\frac{(2j+1)\pi}{2N}v\right] \tag{5}$$

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} & u = 0 \\ \sqrt{\frac{2}{N}} & u \neq 0 \end{cases} \tag{6}$$

$$\alpha(v) = \begin{cases} \sqrt{\frac{1}{N}} & v = 0 \\ \sqrt{\frac{2}{N}} & v \neq 0 \end{cases} \tag{7}$$

where i and j denote pixel index of the input image (size (N, N)); u and v denote DCT coefficient index of the DCT filter. Here, we employed DCT basis functions for one-channel gray images and thus used different bases for each chromatic channel. For example, when using 4 DCT bases, the shape of learnable weight matrix was 16x4x2 (channel number x basis number x chromatic channel), the shape of basis function was 4x1x9x9 (basis number x depth x image width x image height), and the resulted spatial filter had the shape of 16x2x1x9x9. Like for SI-PCA, we varied the number of used basis and picked the one which achieved the best performance on validation data (S1(b) Fig).

Stand-alone EC model (2D). We used a similar EC model architecture (convolutional autoencoder) and loss function as in [39]. The model's encoder contained a single convolutional layer (with weights denoted $\overrightarrow{w_c}$) followed by a rectified linear unit (ReLU) function, one FC layer, and another ReLU function. The decoder contained one FC layer, one ReLU function, a single deconvolutional layer (with weights denoted $\overrightarrow{w_d}$), and a hyperbolic tangent (tanh) function to map back to the original data range ([−1, 1]).

As a measure of reconstruction quality, we used mean squared error (MSE; [38, 39]). We did not use a classical bottleneck with a limited number of units as encoder output layer. Instead, we added Gaussian noise to the encoder output for redundancy reduction [38, 54, 55] and an L1 penalty (hyperparameter β) was imposed to its activation ($\vec{h}$) for sparse readouts [38, 54, 56]. We also applied L2 regularization on the convolutional and deconvolutional layers to encourage the learning of smooth filters [53, 57, 58]. We used 16 9x9 convolutional and deconvolutional filters. The activation tensor (16x28x28, output channel x image width x image height) following the first convolutional layer was flattened to a one-dimensional vector with 12,544 inputs before feeding into the FC layer. The loss function for the EC model was:

$$L_{EC} = \sum_i (\overrightarrow{x_i} - \hat{\overrightarrow{x_i}})^2 + \alpha(\|\overrightarrow{w_c}\|_2 + \|\overrightarrow{w_d}\|_2) + \beta\|\overrightarrow{h}\|_1 \tag{8}$$

where the first term is the MSE error between the prediction $\hat{\overrightarrow{x_i}}$ and ground truth $\overrightarrow{x_i}$ with image index i, and the next two terms denote the L2 and L1 penalties. This way, the EC model learns smooth convolutional filters resembling 2D Gaussians, reminiscent of retinal representations [38, 39].

Hybrid model (2D). The hybrid (semi-supervised) model consisted of a SI and an EC branch (for details on the two models' architectures, see above). These branches were trained simultaneously, sharing the spatial convolutional filters ($\overrightarrow{w_{cs}}$). The total loss function of the hybrid model was derived from the loss functions of the two branches as follows:

$$L_{Hybrid} = wL_{SI} + (1 - w)L_{EC} \tag{9}$$

$$L_{SI} = (\sum_i (\hat{\overrightarrow{r_i}} - \overrightarrow{r_i} log \hat{\overrightarrow{r_i}}) + \alpha_1\|\overrightarrow{w_{cs}}\|_2 + \alpha_2\|\overrightarrow{w_{ct}}\|_2/w + \beta_1\|\overrightarrow{w_f}\|_1/w)/N_1 \tag{10}$$

$$L_{EC} = (\sum_j (\vec{x_j} - \hat{\vec{x_j}})^2 + \alpha_3 \|\vec{w_{cs}}\|_2 + \alpha_3 \|\vec{w_d}\|_2/(1-w) + \beta_2 \|\vec{h}\|_1/(1-w))/N_2 \quad (11)$$

Here, i and j denote neuron and image index, respectively; N_1 and N_2 the number of neurons and images, respectively. The weight (w, with $0 \leq w \leq 1$) controlled the impact of each branch's loss function on the shared spatial filters. Practically, we used $w = 10^{-8}$ for L_{SI} and $w = (1 - 10^{-8})$ for L_{EC} when $w = 0$ and $w = 1$, respectively. Note that we added w to the denominator of the last two terms to maintain the same regularization for $\vec{w_{ct}}$ and $\vec{w_f}$ in a stand-alone SI model when varying w. For L_{EC}, similar to L_{SI}, we added $(1-w)$ to the denominator of the last two terms to keep the same regularization for $\vec{w_d}$ and $\vec{h}$ in a stand-alone EC model when varying w. We used different data to train the EC branch of the hybrid model: natural images, phase-scrambled natural images and noise. All hybrid models were trained for a maximum of 100 epochs (S1(c) and S1(d) Fig); training was stopped early when the prediction on validation data started decreasing.

Tuning all hyperparameters jointly in a grid search was computationally prohibitive. Hence, for the SI branch, we varied the hyperparameters around those determined for the stand-alone configuration ($\alpha_1 = 10$, $\alpha_2 = 10$, $\beta_1 = 1/16$; see above), while for the EC branch, we varied the hyperparameters systematically around the values ($\alpha_3 = 10^3$, $\beta_2 = 1/16$) used in [39]. To tune w, we devised a linear search approach by normalizing the loss functions (using N_1 and N_2). This way, we were able to compare the pure SI and hybrid models, both with best predictive performance, and ensured the performance difference between them came from the EC regularizer.

After training the hybrid model, we estimated the spatio-temporal RFs of all neurons using a gradient ascent algorithm [6]. We visualized the spatial and temporal component of RFs using SVD (cf. Fig 3b), and the magnitude of the RF was indicated in the spatial component.

We trained 2D models using all training data (440 s) with a learning rate of $\mu = 10^{-4}$. In case less data were used (i.e., to evaluate data efficiency), we kept all hyperparameters the same as for the full data case but doubled the learning rate. This was done because the stand-alone SI model and the hybrid model could not reach the minimum of validation loss within 100 epochs (when less data were used).

3D models

Stand-alone SI model (3D). The 3D SI model consisted of one spatio-temporal convolutional layer (16x2x8x9x9, output channels x input channels x depth x image width x image height; depth varied with the frequency of noise stimuli, n = 8 and n = 30 for 5-Hz and 30-Hz noise, respectively), and—after flattening all dimension—one FC layer (96x6,400, output channels x input channels; output channel varied with cell numbers n = 96, 64 or 427 for different data sets; see above), followed by an exponential function. No padding was used. The loss function was defined as:

$$L_{SI} = \sum_i (\hat{\vec{r_i}} - \vec{r_i} log \hat{\vec{r_i}}) + \alpha \|\vec{w_c}\|_2 + \beta \|\vec{w_f}\|_1 \quad (12)$$

This equation differs from Eq (3) with respect to the L2 penalty, which is here on the weights of the spatio-temporal convolutional filters ($\vec{w_c}$) with hyperparameter α for the second term. After performing a grid search for the two hyperparameters, we picked $\alpha = 100$, $\beta = 1/4$ which yielded the best performance on the validation data. After training, we estimated and extracted the cells' spatial and temporal RFs via SVD for visualization.

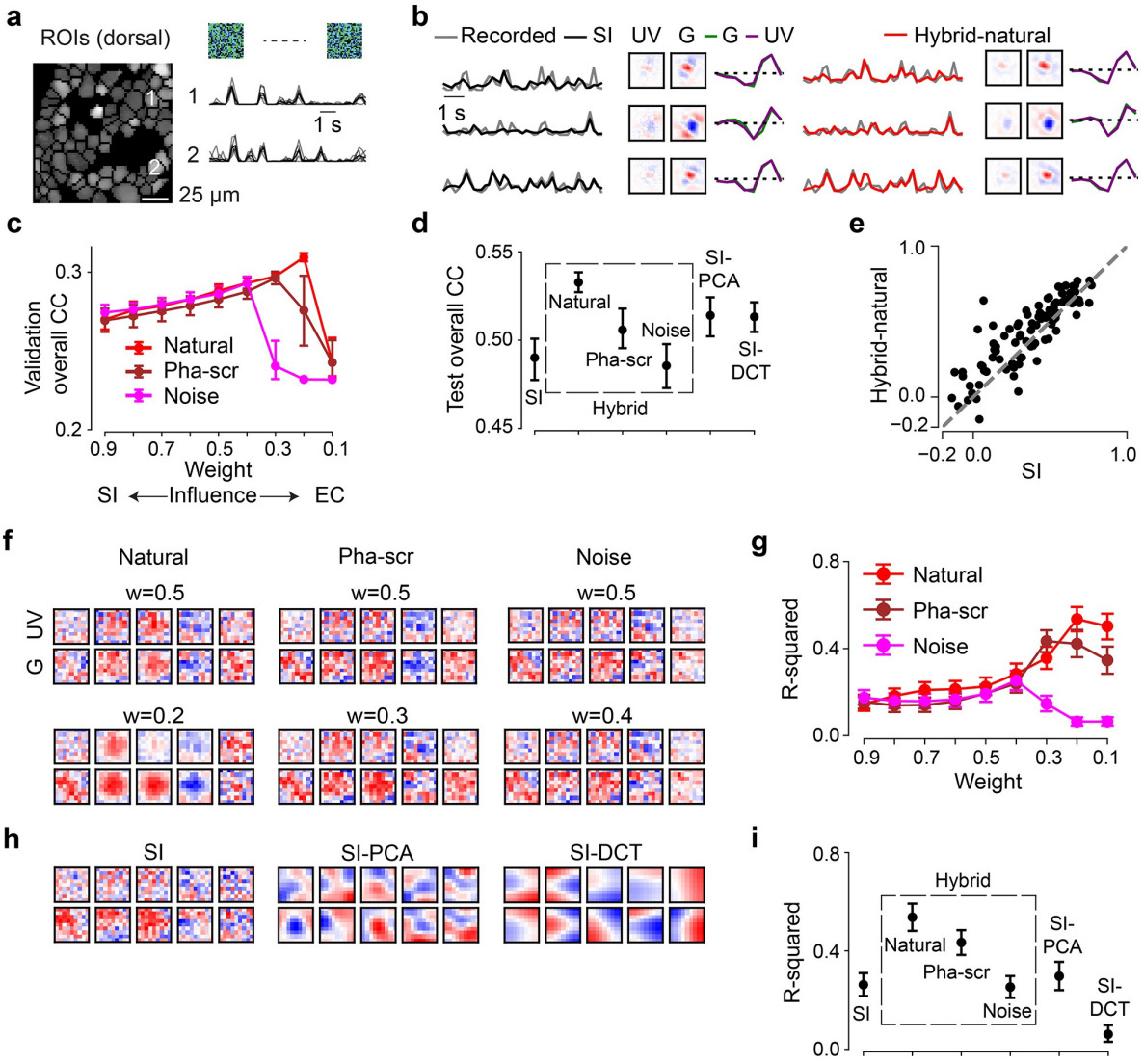

Fig 3. Neural encoding tasks benefit from natural scene statistics. a. Region-of-interest (ROI) mask of one recording field in dorsal retina (left) and mean Ca^{2+} responses (black) of exemplary ROIs in response to 6 repeats of noise stimuli (single trials in gray). **b.** Three representative GCL cell responses (gray) to the noise stimulus (cf. Fig 2a, left), together with predictions of best performing models on test data (black, SI; red, hybrid w/ natural scenes as input to the EC path, i.e., $Input_{EC}$), and learned spatio-temporal receptive fields (RFs) visualized by SVD. **c.** Model performance (linear correlation coefficient, CC; mean for $n = 10$ random seeds per model) based on validation data for hybrid model with natural scenes (red), with phase-scrambled scenes (brown), or with noise (magenta) as $Input_{EC}$, and for different weights. Note that the correlation values for the validation data are relatively low because these predictions were calculated on a single-trial basis (Methods). **d.** Best performance (mean for $n = 10$ random seeds per model) based on test data for SI, *SI-PCA* (16 bases), *SI-DCT* (4 bases), *hybrid-natural* ($w = 0.2$), *hybrid-pha-scr* ($w = 0.3$) and *hybrid-noise* ($w = 0.4$; $p < 0.0001$ for SI vs. *hybrid-natural*, $p = 0.0085$ for *SI-PCA* vs. *hybrid-natural*, $p = 0.0011$ for *hybrid-natural* vs. *hybrid-pha-scr*, two-sided permutation test, $n = 10,000$ repeats). **e.** Scatter plot for model predictions based on test data for *hybrid-natural* ($w = 0.2$) vs. SI at one random seed, with each dot representing one neuron. **f.** Representative spatial filters (shared convolutional filters) for hybrid models with different $Input_{EC}$ and different weights. Upper: with $w = 0.5$; lower: with optimal w (see (c)) for hybrid models. **g.** Mean R-squared of fitting a 2D Gaussian to spatial filters (cf. (f)), for hybrid model with natural scenes (red), with phase-scrambled scenes (brown), or with noise (magenta) as $Input_{EC}$, and for different w (n = 10 random seeds per model). **h.** Representative spatial filters (shared convolutional filters) for SI, SI with PCA filters (16 bases) and SI with DCT filters (4 bases). **i.** Mean R-squared of fitting a 2D Gaussian to the spatial filters for one chromatic stimulus channel (green; $n = 10$ random seeds per model; $p < 0.0001$ for SI vs. *hybrid-natural*, $p < 0.0001$ for *SI-PCA* vs. *hybrid-natural*, $p = 0.0074$ for *hybrid-natural* vs. *hybrid-pha-scr*, two-sided permutation test, $n = 10,000$ repeats). Error bars in (c),(d),(g),(i) represent 2.5 and 97.5 percentiles obtained from bootstrapping.

https://doi.org/10.1371/journal.pcbi.1011037.g003

SI-PCA model (3D). For the 3D SI-PCA models, we applied Eq (4) to the movie clips (2x8x9x9, chromatic channel x depth x image width x image height; depth varied with the frequency of noise stimuli, n = 8 and n = 30 for 5-Hz and 30-Hz noise, respectively). Like for 2D SI-PCA models, we varied the number of used bases and picked the number for which the model achieved the best performance on the validation data (S3(a) Fig).

Stand-alone EC model (3D). The 3D EC models used a sequence of frames from a movie clip as input and featured 3D spatio-temporal convolutional layers (with weights denoted $\overrightarrow{w_c}$) in the encoder. The decoder contained deconvolutional layers with weights $\overrightarrow{w_d}$. In the past-encoding case, we fed an 8-frame clip (frames at $t - 7$ to t) to the model and aimed at reconstructing the 7th frame (at $t - 1$). In the future-prediction case, the goal was to predict the 8th frame (at t) with the input being the first 7 frames ($t - 7$ to $t - 1$) of the clip. The loss functions was similar to that given by Eq (8) except that (*i*) $\overrightarrow{w_c}$ features different a shape (16x2x8x9x9, output channel x chromatic channel x filter depth x filter width x filter height), and (*ii*) x_i denotes the 7th frame for the past encoding and the 8th frame for the future prediction model (S2(b), S2(c) and S2(d) Fig).

Hybrid model (3D). The 3D hybrid models consisted of a SI branch and an EC branch with shared spatio-temporal convolutional filters ($\overrightarrow{w_c}$; see above). Like for the 2D hybrid models, the total loss function was a weighted sum of losses for the two branches as follows:

$$L_{Hybrid} = wL_{SI} + (1 - w)L_{EC} \tag{13}$$

$$L_{SI} = (\sum_i (\hat{\overrightarrow{r_i}} - \overrightarrow{r_i} log \hat{\overrightarrow{r_i}}) + \alpha_1 \|\overrightarrow{w_c}\|_2 + \beta_1 \|\overrightarrow{w_f}\|_1 / w)/N_1 \tag{14}$$

$$L_{EC} = (\sum_j (\overrightarrow{x_j} - \hat{\overrightarrow{x_j}})^2 + \alpha_2 \|\overrightarrow{w_c}\|_2 + \alpha_2 \|\overrightarrow{w_d}\|_2/(1 - w) + \beta_2 \|\overrightarrow{h}\|_1/(1 - w))/N_2 \tag{15}$$

Here, i denotes neuron index, j movie clip index, N_1 neuron number, and N_2 the number of movie clips. Again, instead of tuning all hyperparameters jointly via a grid search, we varied the hyperparameters around the values determined for the stand-alone SI configuration (α_1 = 100, β_1 = 1/4) for the SI branch. For the EC branch, we varied the hyperparameters systematically around the values ($\alpha_2 = 10^4$, β_2 = 1/16) used in the stand-alone EC models. We then tuned w linearly after normalizing the loss functions (using N_1 and N_2). We also visualized the spatial and temporal RF components using SVD.

Results

Hybrid system identification and efficient coding models

To test if learning an efficient representation of natural input could help predict neuronal responses in the early visual system, we employed *normative regularization*, i.e., statistical regularization that is informed by normative coding principles, such as the idea that sensory systems have evolved to efficiently process natural stimuli. Specifically, we used this strategy to incorporate EC as a regularizer and developed a hybrid model that combines SI-based neural prediction and EC in a single model. The two model branches are linked by shared convolutional filters (Fig 1b).

The *SI branch* approximates the response functions of recorded neurons to a visual dense noise (see below), and was implemented using a convolutional neural network (CNN) (Fig 2a). Here, we used an L2 regularization on the convolutional layers to encourage smooth filters

[53] and an L1 regularization on the fully connected (FC) layer for sparse readouts ([19]; for details, see Methods).

The *EC branch* was trained to efficiently reconstruct input stimuli (i.e., natural scenes) from a constrained latent representation. For this branch, we used a convolutional autoencoder network that we published before (for details, see [39] and Methods). Also in the EC branch, we enforced smooth filters by using L2 regularization. In addition, we limited the bandwidth by adding Gaussian noise and imposing L1 regularization on the hidden activations. The latter regularization also encourages sparse representations.

In the *hybrid model*, we implemented interactions between the two branches by shared filters (symbolized by red circle in Fig 1b). Both branches were trained in parallel, with a weighted sum of their respective losses (L_{SI} and L_{EC}) used as optimization objective. By changing the weighting of the two losses, we were able to control the relative contribution of two branches on shaping the shared filters, and test our hypothesis to which degree efficient representations of natural scenes improve neural predictions (Fig 2a and 2b). Specifically, weight w was used to define the hybrid model's loss function as $L_{Hybrid} = w \cdot L_{SI} + (1 - w) \cdot L_{EC}$ (Methods). For $w = 1$, the EC branch had no influence on the shared filters and, hence, the hybrid model behaved like the pure SI model. Conversely, for $w = 0$, the SI branch had no influence on the shared filters and, hence, the hybrid model behaved like the pure EC model. Thus, the smaller the weight, the more the EC branch contributed to shaping the filters.

To evaluate the influence of stimulus statistics on neural response predictions, we fed not only natural stimuli to the EC branch, but also phase-scrambled natural stimuli as well as noise. We refer to these models as *hybrid-natural*, *hybrid-pha-scr* and *hybrid-noise* (Fig 2c). Moreover, to examine whether the performance improvements could be attributed to simple low-pass filtering, we trained SI networks using spatial convolutional filters composed of different numbers of basis functions derived from principle component analysis (PCA) on natural images (Fig 2d), or the discrete cosine transform (DCT). These models are referred to as *SI-PCA* and *SI-DCT* networks.

To train the SI branch of our hybrid framework, we recorded somatic Ca^{2+} responses from populations of cells in the ganglion cell layer (GCL) of the *ex-vivo* mouse retina to 9-minute long noise stimuli using two-photon imaging (Fig 3a; Methods; [44, 47]). The GCL contains the RGCs, which represent the retina's output neurons and form in the mouse about 40 parallel feature channels to higher visual brain areas (reviewed in [23]). RGCs gain their specific response properties by integrating upstream input from distinct sets of bipolar cells and amacrine cells. Note that the GCL also contains some "displaced" amacrine cells (dACs; [47, 59]). If not indicated otherwise, we did not distinguish between these two GCL cell classes in our datasets. The noise stimulus contained two chromatic components (UV, green) matching the spectral sensitivities of mouse photoreceptors [60]. We used the data of n = 96 GCL cells that passed our quality criteria (Methods) to fit a pure SI model with factorized spatial and temporal convolutional filters, whose predictive performance served as our baseline (Fig 3b left).

Neural system identification benefits from natural scene statistics

First, we measured the predictive performance of the *hybrid-natural* model on the validation data (for hyperparameter tuning) by systematically varying the relative impact of the two branches, i.e., changing the weight w. We found that the performance steadily increased with increasing EC influence (i.e., decreasing w) up to an optimum (peaking at $w = 0.2$; Fig 3c, red), after which the SI had too little influence on the shared filters and the performance dropped.

Next, we replaced the natural input to the EC pathway by phase-scrambled scenes (*hybrid-pha-scr*) and white noise across space and chromatic channels (*hybrid-noise*). Like for the

hybrid-natural model, the performance of the two control models also increased with increasing EC influence up to a certain point, peaking at $w = 0.3$ and $w = 0.4$ for *hybrid-pha-scr* and *hybrid-noise*, respectively (Fig 3c). This indicates that when incorporating EC, all hybrid model versions showed some improvement up to certain w values, before performance sharply declined.

To test to what extent simple low-pass filtering contributes to the performance improvement observed for the *hybrid-natural* model, we quantified the performance of two additional SI models, one with PCA and the other one with DCT bases. By varying the number of bases used, we found a maximum in predictive performance at 16 and 4 bases for *SI-PCA* and *SI-DCT* (zig-zag ordering), respectively (S1(b) Fig).

Finally, to compare the performance on the test data across models, we picked for each model the w or number of bases with the best predictive performance for the validation data. We found that the hybrid model with natural inputs to the EC branch attained the best performance among all tested models (Fig 3d and 3e). The *hybrid-natural* model's superior performance compared to the *hybrid-pha-scr* model suggests that the benefit of learning natural scene statistics extends beyond second-order statistics such as the $1/f$ power spectrum of natural images. Nevertheless, the *hybrid-pha-scr* model performed better than the *hybrid-noise* version, pointing at a general benefit of learning second-order statistics in the EC branch. Moreover, the *hybrid-natural* model was consistently better than low-pass filtering control models (*SI-PCA* and *SI-DCT*), suggesting that simple low-pass filtering does not fully explain the benefits of sharing kernels with the EC branch trained to efficiently represent natural stimuli.

Together, our results suggest that normative network regularization—in particular, based on natural statistics—can improve the performance of neural SI models on predicting responses to noise.

Hybrid models with natural inputs learn the most "biologically-plausible" filters

To confirm that our hybrid models capture the properties of the recorded cells, we estimated their RFs (Fig 3b and S1(f) Fig; Methods). Indeed, we found that the models learned antagonistic center-surround RFs with biphasic temporal kernels, reminiscent of RGC RFs found in other studies [2, 47]. To get insights to which degree our models resembled biological vision systems, we next investigated the internal representations by analyzing the filters of the models' subunits [18, 61]. To this end, we compared the shared spatial convolutional filters between our tested models. As neurons in the retina and further upstream in the early visual system often feature smooth, Gaussian or DoG shaped RFs (e.g., [47, 62, 63]), we refer in the following to models with such shared filters as more "biological plausible" than those with other filter organizations.

Interestingly, while the learned neuronal RFs were quite consistent between models (cf. Fig 3b), their shared spatial filters differed considerably (Fig 3f and 3h). When using natural images in the EC branch (*hybrid-natural*), filters indeed became smoother and more Gaussian-shaped, which may be due to the regularization by the EC branch on the SI branch and may have contributed to the performance improvement of predicting responses. This effect persisted though reduced when phase-scrambled images were used (*hybrid-pha-scr*). Moreover, for smaller w values (i.e., stronger EC influence), Gaussian-shaped filters became more frequent in the *hybrid-natural* but not in the *hybrid-noise* model (Fig 3f, upper vs. lower row). For the SI models with PCA or DCT basis, we found all filters to be smooth as they profited

from low-pass filtering of the respective transformation. However, compared to the *hybrid-natural* model, their filters were less frequently Gaussian-shaped (Fig 3h).

To quantify these findings, we fit 2D Gaussian functions to the filters and measured the goodness of the fit via the coefficient of determination (R-squared; Methods). Notably, for all three hybrid models, the w with the best Gaussian fit was the same w that also resulted in the best response predictive performance ($w = 0.2$, $w = 0.3$, and $w = 0.4$ for *hybrid-natural*, *hybrid-pha-scr*, and *hybrid-noise*, respectively; Fig 3g). The filters of the *hybrid-natural* model resembled smooth 2D Gaussians more than for any other model (Fig 3i), including *SI-PCA* and *SI-DCT*. The difference of fit quality between *hybrid-natural* vs. *hybrid-pha-scr* and *hybrid-pha-scr* vs. *hybrid-noise* may be related to higher-order statistics and second-order statistics of natural scenes, respectively.

Taken together, our comparisons of the hidden spatial representations suggest that natural scene statistics promote latent feature representations akin to transformations in the early visual system.

Efficient coding increases the data efficiency of system identification

Next, we asked if the observed performance increase in the *hybrid-natural* vs. the baseline SI model was sensitive to the amount of training data, both with respect to their response predictions (Fig 4a) and their learned spatial filters (Fig 4b). To this end, we trained the SI and the *hybrid-natural* model ($w = 0.2$) with different amounts of data, ranging from 30% to 100%.

Not unexpectedly, when more training data was used, predictive performance increased for both models (Fig 4a top). However, we also found that the performance of the *hybrid-natural* model was consistently higher than that of the SI model, with the difference becoming significant for $\geq$ 60% and peaking at around 90% training data (Fig 4a bottom). Additionally, for both models the spatial filters became increasingly more Gaussian-like with more data (Fig 4b). We also observed that the performance difference dropped for large dataset sizes—which, we expect, is asymptotically near zero in the regime of infinite data.

Together, these results suggest that a hybrid model predicting responses to noise, but with access to natural statistics requires significantly less training data than the baseline SI model.

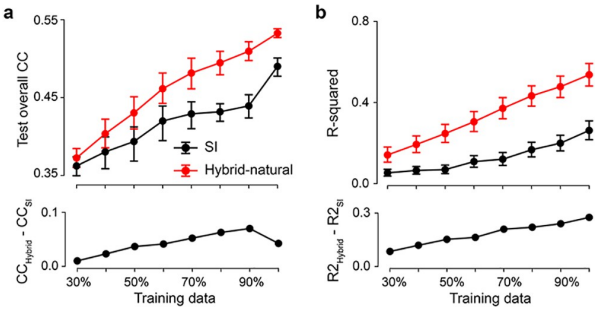

Fig 4. *Hybrid-natural* models have better data efficiency for neural prediction. a. Mean model performance (top) based on test data for SI and *hybrid-natural* ($w = 0.2$; $n = 10$ random seeds) with different training data sizes and mean difference between SI and *hybrid-natural* (bottom). **b.** Mean R-squared (top) of fitting a 2D Gaussian to spatial filters for green stimulus channel for SI and *hybrid-natural* ($w = 0.2$; $n = 10$ random seeds) with different training data sizes, and the mean difference between R-squared for SI and *hybrid-natural* (bottom). Error bars represent 2.5 and 97.5 percentiles with bootstrapping.

https://doi.org/10.1371/journal.pcbi.1011037.g004

Hybrid models for testing temporal coding strategies

It has been suggested that early stages of visual processing, rather than encoding a past stimulus (past encoding), aim at predicting future stimuli in their temporal stream of inputs [64]. Such a future prediction strategy is thought to extract information that can be used for guiding future behavior [65]. Therefore, we next tested if sharing spatio-temporal (i.e., 3D) filters can further improve the predictive performance of the hybrid model. We implemented both strategies—past encoding and future prediction—in the EC branch, and compared their influence on the SI task [66].

We modified the 2D SI model to use spatio-temporal (instead of factorized spatial and temporal) convolutional filters to predict neural responses for 8-frame noise movies (3D SI model; S2(a) Fig). Likewise, we employed spatio-temporal convolutional filters for the EC branch. As before, the two branches of the resulting hybrid model were trained in parallel, but now sharing spatio-temporal filters. In the past encoding case, the EC branch was trained to reconstruct the 7^{th} frame (at $t - 1$) of a continuous 8-frame natural movie clip based on frames at $t - 7$ to t (*hybrid-natural-past*; S2(b) and S2(c) Fig). In the future prediction case, the EC branch was trained to predict the 8^{th} unseen frame based on the first 7 frames ($t - 7$ to $t - 1$) of the clip (*hybrid-natural-future*; S2(d) Fig left).

Like for the 2D models, we varied w or the number of bases and then selected the best model for each condition (3D SI, *hybrid-natural-past*, *hybrid-natural-future*, and 3D *SI-PCA*) based on validation performance. We next quantitatively compared the different models using the test data (Fig 5a and 5b; S3c Fig). We found that the 3D *SI-PCA* model outperformed the 3D SI model, presumably because the former profited from the low-pass filtering of the PCA transformation. Importantly, both hybrid models displayed a better performance than the 3D *SI-PCA* model. While the *hybrid-natural-past* model performed slightly better than its *hybrid-natural-future* counterpart, this difference was not statistically significant. In summary, both the past encoding and future prediction strategy in the EC branch turned out to be equally beneficial for predicting the responses to noise stimuli and, as before, the benefit extended beyond low-pass filtering effects. However, no performance increase was achieved with respect to the 2D *hybrid-natural* model (Fig 5b vs. Fig 3d).

We also analyzed the shared spatio-temporal filters using the same metric as for the 2D case, which assesses the similarity between spatial filters (after performing a low-rank decomposition of 3D shared filters into spatial and temporal components; see Methods) and smooth 2D Gaussians (Fig 5c and 5d). Again, we found higher R-squared values for the hybrid models and the 3D *SI-PCA* model compared to the baseline SI case. Note that here, the 3D *SI-PCA* model did not significantly differ from the two hybrid models, possibly due to a large number of bases ($n = 128$ vs. $n = 16$ in the 2D case).

Next, we asked if the fact that we did not see a significant advantage of 3D over 2D could be because the slow (5 Hz) noise stimulus did not sufficiently drive GCL cell responses. Therefore, we recorded a dataset ($n = 64$ cells), in which we presented a 30-Hz dense noise stimulus and used it with the 3D hybrid models. Like for 5-Hz noise, the *hybrid-natural-past* and *hybrid-natural-future* models performed better than the 3D SI model, both on response prediction and with higher R-squared values for the learned filters (S4 Fig). But again, the 3D hybrid models performed only equally well compared to their 2D counterparts.

In summary, the *hybrid-natural* models achieved a higher performance for different noise stimuli (5-Hz vs. 30-Hz) and different shared filter organizations (2D vs. 3D) than all other tested models. Therefore, it is likely that their superior predictive performance for neuronal responses and their more biologically plausible filters resulted from the EC branch having access to natural statistics.

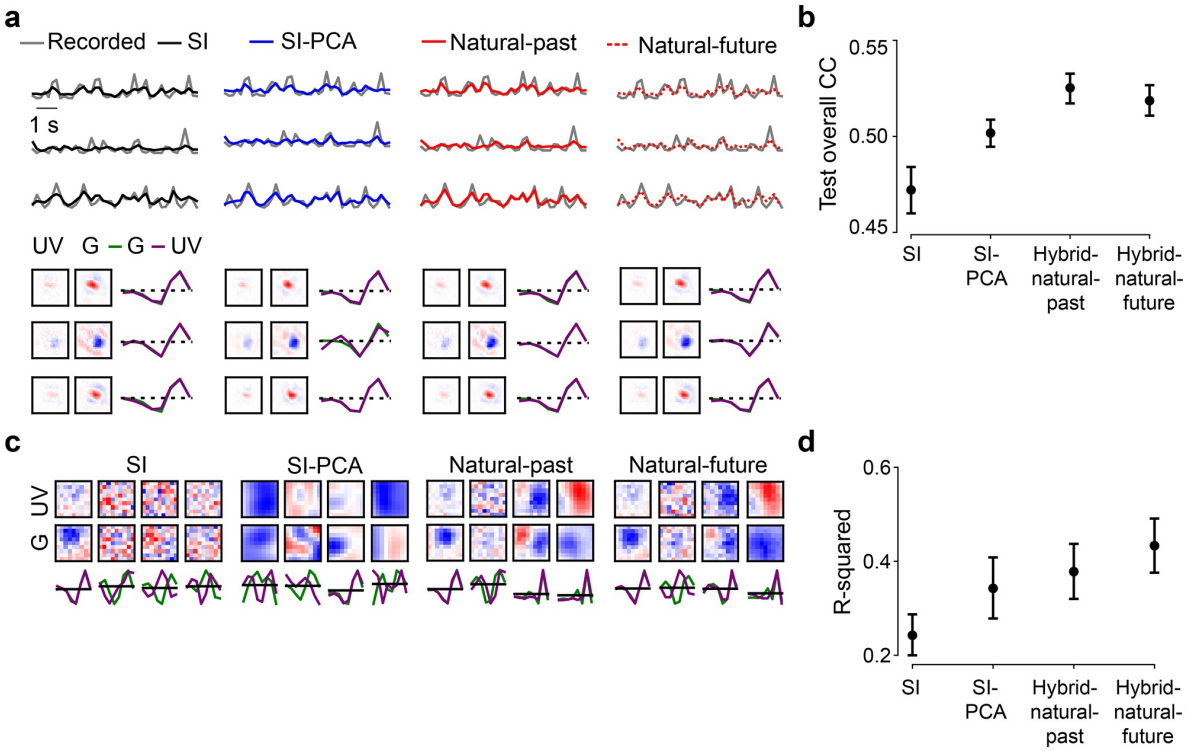

Fig 5. Past encoding or future prediction strategies using 3D shared filters perform equally well. a. Top row: Responses of three exemplary GCL cells to 5-Hz noise stimulus (gray) and predictions of best performing models on test data (black, SI; blue, SI with PCA filters; red solid, hybrid for encoding the past; red dotted, hybrid for predicting the future). Bottom row: Respective learned RFs of the three cells (visualized by SVD). **b.** Mean model performance based on test data for SI, SI-PCA (128 bases), *hybrid-natural-past*, and *hybrid-natural-future* (both $w = 0.4$; $n = 10$ random seeds; $p < 0.0001$ for SI vs. *hybrid-natural-past*, $p = 0.0005$ for SI-PCA vs. *hybrid-natural-past*, $p = 0.2563$ for *hybrid-natural-past* vs. *hybrid-natural-future*, two-sided permutation test, $n = 10,000$ repeats). **c.** Representative shared spatial and temporal filters of 3D models (n = 1 random seed, visualized by SVD; temporal kernels for UV and green stimulus channels indicated by purple and green, respectively). **d.** Mean R-squared of fitting a 2D Gaussian to shared spatial filters (for green stimulus channel; $n = 10$ random seeds per model; $p = 0.0003$ for SI vs. *hybrid-natural-past*, $p = 0.4356$ for *SI-PCA* vs. *hybrid-natural-past*, $p = 0.1895$ for *hybrid-natural-past* vs. *hybrid-natural-future*, two-sided permutation test, n = 10,000 repeats). Error bars in (b),(d) represent 2.5 and 97.5 percentiles with bootstrapping.

https://doi.org/10.1371/journal.pcbi.1011037.g005

Direction-selective neurons benefit more than others from hybrid models

The retina encodes the visual scene in a number of features that are represented by more than 40 different types of RGC whose outputs are relayed in parallel to higher visual centers in the brain [47, 67–70]. Thus, we next asked, if access to natural statistics allowed our hybrid models to predict some cell types better than others (Fig 6). Earlier, it has been shown that motion-relevant properties emerge in the efficient coding framework for both past encoding and future prediction approaches [66]. Therefore, we employed our 3D hybrid models (cf. Fig 5) and focused on direction-selective (DS) cells [47, 71].

For this analysis, we used a set of $n = 427$ GCL neurons, whose responses were recorded not only to the 5-Hz noise stimulus (for training the models) but also to full-field chirp and moving bar stimuli. The latter two stimuli (Fig 6a) enabled us to identify the functional type of each recorded GCL neuron [47] using a cell type classifier (see Methods; S5 Fig).

We observed that for 100% of the data, SI and hybrid model performed similarly well. For the analysis of cell type-specific performance, we therefore chose a dataset size (30% of total

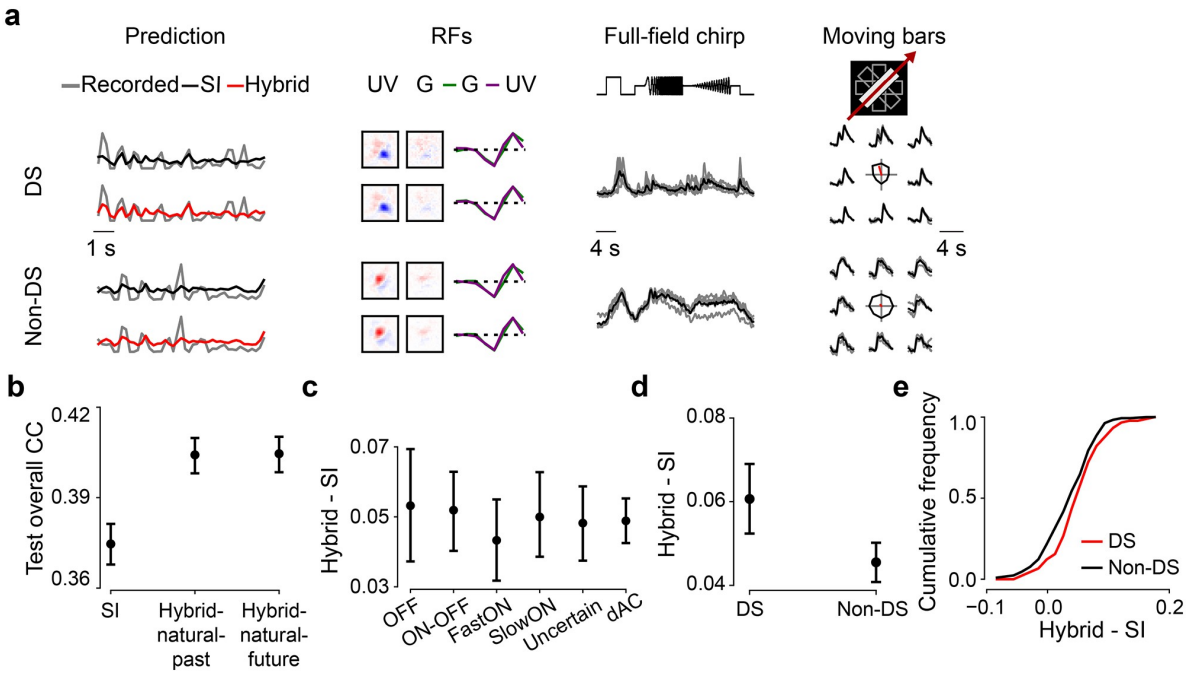

Fig 6. Direction-selective (DS) neurons benefit more from hybrid models. a. Recorded (gray) and predicted (black, SI; red, *hybrid-natural-past*; response amplitude scaled with a constant 1.5 for better visualization) responses to noise, RFs, as well as full-field chirp responses and moving bar responses (gray, single trials; black, means) of representative DS and non-DS cells. Note that the RFs were dominated by UV stimulus channel because cells were recorded in ventral retina (see Methods). **b.** Mean model performance based on test data for SI, *hybrid-natural-past* and *hybrid-natural-future* (both $w = 0.7$; $n = 10$ random seeds per model; trained with responses of $n = 427$ GCL cells to 5-Hz noise stimulus; $p < 0.0001$ for SI vs. *hybrid-natural-past*, $p = 0.9307$ for *hybrid-natural-past* vs. *hybrid-natural-future*; two-sided permutation test, $n = 10,000$ repeats). Note that compared to Fig 5b, these models had a lower predictive performance, as we used a different dataset, with 30% of data for training. **c.** Difference in mean performance between *hybrid-natural-past* and SI based on test data for 6 broad functional groups of GCL cells (35 OFF, 59 ON-OFF, 49 fast-ON, 38 slow-ON, and 64 uncertain RGCs, as well as 145 dACs; see Methods and Results; $n = 10$ random seeds per model). **d.** Like (b) but for $n = 90$ DS and $n = 300$ non-DS cells. **e.** Cumulative histogram of difference in mean prediction between *hybrid-natural-past* ($w = 0.7$) and SI on test data for DS (red) and non-DS cells (black), at one particular seed. Error bars in (b)–(d) represent 2.5 and 97.5 percentiles with bootstrapping.

https://doi.org/10.1371/journal.pcbi.1011037.g006

recording time) for which the predictive performance difference between the two models was particularly pronounced. As expected, we found that both hybrid networks (*hybrid-natural-past* and *hybrid-natural-future*) performed significantly better than the SI model, with no significant difference between the two hybrid models (cf. Fig 5b, S4(b) Fig).

First, we evaluated if any of the broader functional groups of GCL cells profited more from natural statistics than others. For this, we sorted the cells into 6 groups based on their response polarity (ON vs. OFF) and transience, and based on whether they were RGCs or dACs (for group sizes, see Fig 6 legend). For all 6 groups, the hybrid models showed a better predictive performance than the SI model (Fig 6b). However, no significant differences were observed between any pair of groups ($p > 0.05$ for all pair-wise comparisons, two-sided permutation test, n = 10,000 repeats; Fig 6c) and the two hybrid models ($p > 0.05$ for all pair-wise comparisons; S6(a) Fig).

Next, we grouped the cells into DS ($p < 0.05$, direction tuning using a permutation test; $n = 90$) and non-DS cells ($n = 300$) based on their moving bar responses (Fig 6a right). Note that $n = 37$ neurons were excluded as they did not pass the quality test for chirp and moving-bar responses (Methods). We found that the predictive performance for DS cells was significantly higher than that of the non-DS cells for both *hybrid-natural-past* (Fig 6d and 6e;

$p = 0.0027$) and *hybrid-natural-future* (S6(b) and S6(c) Fig; $p = 0.0042$). To test whether this performance difference was merely due to different signal-to-noise ratios in DS vs. non-DS cells, we compared their response quality indices (*QI*; Methods). While DS cells had significantly higher *QI* values for moving-bar responses (QI_{bar}) than non-DS cells, we did not find any significant difference between the two groups with respect to their noise (QI_{noise}) or chirp responses (QI_{chirp}; S6(e), S6(f) and S6(g) Fig). These results suggest that DS cells benefit more from the EC branch of the hybrid models than non-DS cells, partially consistent with earlier findings ([66]; see also Discussion).

In summary, efficient coding of natural statistics served as a beneficial normative regularization for predicting responses to noise stimuli of all types of mouse GCL cells and in particular DS cells, suggesting the potential role of motion statistics in the natural environment on shaping neuronal response properties.

Hybrid models for predicting retinal responses to natural movies

Natural stimuli are thought to drive more diverse neural responses compared to artificial stimuli, such as dense noise or drifting bars [72]. As a result, more complex feature transformations are expected to be required for determining the respective stimulus-response functions ([18, 73], but also see [74]). Therefore, we tested if predicting neural responses to natural movies would also profit from our hybrid model.

To this end, we used the neural activity of $n = 86$ ventral GCL neurons that were presented with 30-Hz natural movies ([39]; Fig 7a left) to train a stand-alone SI model with factorized spatial and temporal filters. Surprisingly, the SI model learned center-surround RFs with biphasic temporal components as well as smooth 2D Gaussian spatial filters with high R-squared values (mean R-squared = 0.96; full training data of 433 s; Figs 7a, 3f, 3h and 4b).

Therefore, we next tested if there was a performance difference between the SI and hybrid models for less training data (Fig 7b and 7c). Here, we decided to use approx. a quarter of the data (i.e., 23%, or 15 of 65 mini-batches). As with the models using neural responses to noise, we tuned the hyperparameters based on validation data ($w = 0.4$ for *hybrid-natural*) and fixed them. We then evaluated the performance of the two models after being trained with different

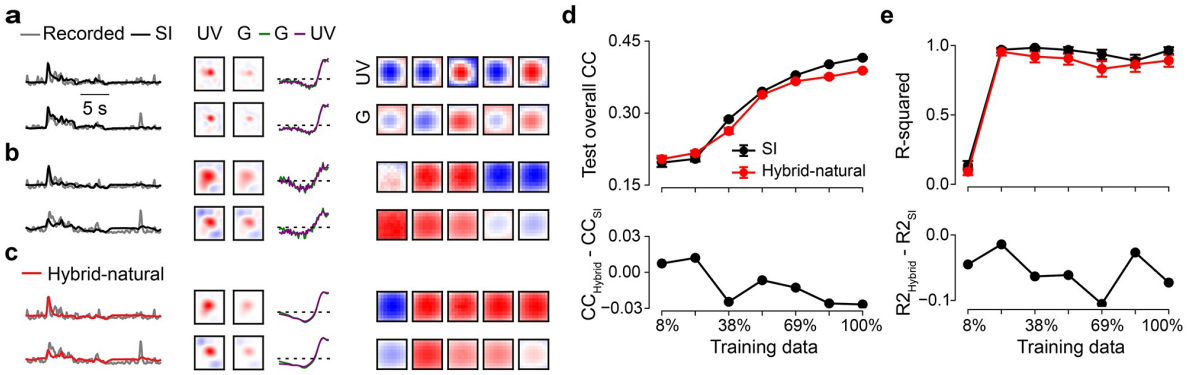

Fig 7. Predicting neural responses to natural movies does not benefit from efficient coding. a. Recorded (gray) and predicted (black, SI; red, *hybrid-natural*) responses to natural movie, RFs, as well as exemplary spatial filters for the SI model trained by full training data. **b.** Same as (a), but for the SI model trained by 23% of training data. **c.** Same as (a), but for the *hybrid-natural* model trained by 23% of training data. **d.** Predictive performance (top) based on test data for SI and *hybrid-natural* ($w = 0.4$; $n = 10$ random seeds) with different training data sizes, and the difference between SI and *hybrid-natural* (bottom). **e.** R-squared (top) of fitting a 2D Gaussian to spatial filters for UV stimulus channel for SI and *hybrid-natural* ($w = 0.4$; $n = 10$ random seeds) with different training data sizes, and the difference between SI and *hybrid-natural* (bottom). Error bars in (d),(e) represent 2.5 and 97.5 percentiles with bootstrapping.

https://doi.org/10.1371/journal.pcbi.1011037.g007

amounts of data, ranging from 8% to 100%. As expected, predictive performance of both models increased with available data. However, compared to the models trained with noise responses, the *hybrid-natural* model had similar predictive performance with the SI model, with only a marginal improvement for small amounts of data (<25%) (Fig 7d; cf. Fig 4a; see Discussion).

Additionally, the two models had similar R-squared values across different data sizes (*hybrid-natural* model with slightly lower values; Fig 7e), indicating that they learned filters that resembled 2D Gaussians comparably well. We also observed that both models featured filters with high R-squared values starting from 23% of the data (cf. Fig 4b).

Together, these results suggest that normative network regularization, as implemented in our hybrid models, offers no additional benefit for predicting responses to natural movies.

Discussion

In this study, we asked if access to natural scene statistics can help predicting neural responses. To address this question, we combined system identification (SI, [3]) and efficient encoding (EC, [25]) methods into a normatively regularized (hybrid) modeling framework. Specifically, we used models that efficiently represent natural scenes recorded in the mouse' habitat to regularize models that predict retinal responses to visual stimuli. We analyzed such hybrid models with shared spatial filters, and found that natural images as input to the EC branch indeed improved the performance in predicting retinal responses to noise stimuli and allowed the model to generate filters that resembled RFs found in the early visual system. These improvements extend beyond those gained by simple low-pass filtering or using second-order statistics of the natural scenes. Our hybrid models with shared spatio-temporal filters performed similarly well as those with shared spatial filters, independently of whether they used a past encoding or a future prediction strategy. Notably, predictions of DS cell responses to noise stimuli improved the most in the hybrid models with natural input to the EC branch. Interestingly, in predicting neural responses to natural movies, both hybrid and SI models performed similarly well. In summary, our results suggest that sourcing information about an animal's environment—e.g., through hybrid SI-EC models—can help building more predictive and biologically-plausible models of neuronal networks—at least when predicting neural responses to artificial stimuli and/or for limited amounts of data. More generally, our findings lend support to the idea that knowledge of natural statistics is already encoded in sensory circuits.

Hybrid models improve data efficiency

When predicting responses to noise, the difference in predictive performance between the hybrid and the baseline SI model was significant and it depended on the amount of available data, indicating that our hybrid modeling approach increased data efficiency. The data efficiency also depended on the input to the SI branch in the hybrid models: For natural stimuli, the performance gain was marginal and restricted to the case when data was strongly limited. Therefore, we expected our hybrid models to improve SI mainly when only little neural data in response to artificial stimuli is available. It is possible that for those more challenging problems at downstream visual areas, where neural response functions and, hence, the neural prediction tasks, become more complex [75], the data efficiency of a hybrid approach and the improvement from natural scene statistics may be higher.

Biological plausibility and temporal coding principles in hybrid models

Regarding the spatial filters, for most learned models the degree of similarity to Gaussian RFs was positively correlated with their predictive performance (with the exception of the *SI-DCT*

models)—whether there is a causal link remains unclear (see below). Note that we used the filters' similarity to 2D Gaussian functions as a proxy for biological plausibility, following the assumption that for RFs in the retina and at early downstream stages of the visual system, a smooth, Gaussian-like structure is often a suitable approximation (e.g. [47, 62, 63]). Accordingly, the fitted Gaussian shapes had diameters of 4–9 pixels, equivalent to 3.3˚-7.4˚ of visual angle and, hence, in the range of RF center sizes of mouse RGCs (3˚-13˚; [76–78]). However, it has been reported that RGCs, for instance, can also feature multiple sensitivity peaks and irregular non-Gaussian shapes in their RF (e.g., [79]). Therefore, our proxy of biological plausibility may underestimate the complexity of retinal representations and for future studies, it would be important to use additional filter properties, such as locality and smoothness, as metrics. Moreover, a deep, systematic understanding of artificial and neuronal networks and their hidden representations likely calls for other methods besides filter inspection, e.g., the evaluation of temporal curvature ([80, 81]; discussed in [82]). As the natural environment is not static, we also created hybrid models that acknowledged the time domain by sharing spatio-temporal filters. Surprisingly, both variants—past encoding and future prediction—behaved quite similarly. Note that our future prediction approach is not the same as "predictive coding", which removes redundancy in uniform or correlated inputs by encoding the difference between the actual input and the internal expectation [24, 83–85]. However, in the stand-alone EC models (that is, only the EC branch), the temporal components of the filters learned by the future prediction were much more diverse than those of past encoding (S2(c) and S2(d) Fig right). Interestingly, the differences between temporal filters of these stand-alone EC models decreased with the incorporation of the neural prediction task in the hybrid models.

The filter diversity in our 3D hybrid models is reminiscent of earlier findings by Chalk and colleagues [66], who reported the emergence of filters sensitive to motion direction and motion speed in their past encoding and future prediction EC models, respectively. However, in contrast to their results, we did not see a difference between our *hybrid-past* and *hybrid-future* models with respect to motion-sensitive filters: Both of them performed better in predicting responses of DS vs. non-DS cells. Further work is needed to understand that partial (mis)match between our work and that by Chalk et al., and why specifically DS cells profited from both our 3D hybrid models. It is possible that the better performance for DS cells is related to the fact that the natural movies we used for training the EC branch are dominated by global motion [39]. In other words, our EC model may be prone to produce filters that detect spatio-temporal structures inherent in the training input.

Hybrid models of retinal signal processing

Only for very limited data, our hybrid models displayed a marginal improvement when predicting neural activity to the natural movie stimulus vs. the noise stimulus. This was surprising, as we expected that the EC branch supports the learning of complex feature representations driven by natural stimuli [18, 73]. That the stand-alone SI model trained with natural movie responses easily learned smooth Gaussian filters may have limited the benefits from the hybrid model. In turn, this may indicate that, indeed, predictive model performance correlated with biological filter plausibility. To further explore the interaction between learned filters and predictive performance, it may be instructive to test a greater variety of stimuli and record responses to them from the same neurons. Such data may also be useful for characterizing model generalization (domain transfer, see e.g., [73, 86]) by using responses to natural stimuli as unseen test data with a hybrid model trained with cell responses to noise stimuli. Here, one would need to take into account that RGCs may adapt to different stimuli (such as noise vs. natural movies) by changing their RF properties [84].

Many studies have applied EC principles to natural images with different regularization strategies, such as adding noise (to the input, hidden activation or output), forcing sparsity (of weights, hidden activation or responses), and encouraging smoothness and spatial locality of weights. These coding principles produced diverse feature representations, including DoG and Gabor filters [38, 39, 87–89]. For example, Doi et al. [89] found that the response sparsity and the spatial locality of filter weights induced oriented and center-surround structures, respectively. In our previous study [39], we tested a convolutional autoencoder using a smoothness constraint (L2 regularization) on the convolutional/deconvolutional filters, and a sparsity constraint (L1 regularization) as well as Gaussian noise on the encoder output. We found that this model produced center-surround filters when trained with natural images. Similarly, Ocko et al. [38] trained an autoencoder model with pink noise and obtained DoG filters. Our study shows that *hybrid-natural* had a better predictive performance and biological plausibility than *hybrid-pha-scr*. In turn, *hybrid-pha-scr* outperformed *hybrid-noise*. This suggests that both the regularization and the statistical properties of the model input we used contributed to the emergence of center-surround features.

Generally, the effect of normative network regularization depends on many factors, including—in our hybrid models—neural prediction tasks (e.g., predicting responses to noise vs. natural movies), normative principles (e.g., encoding the past vs. predicting the future), stimulus input of the EC branch (e.g., noise vs. natural scenes), and shared components between two branches (e.g., filter weights vs. network features). Any of them may influence the model performance or the learned filter representations. A factor that we did not vary much was nonlinearity of the model. For example, that our hybrid approach did not improve the prediction of responses to natural movies, which are highly non-linear and complex [18, 73], may be due to limited expressive power of the EC network and shared units.

For our current analysis, we used broad group assignments (e.g., Fast ON RGCs), which include several functional types of RGC (e.g., ON-step, ON-transient, ON-high-frequency etc; [47]) or dACs, but did not detect any differences in performance gain except for the DS neurons. Still, it is possible that distinct types of RGC profit more than others from the EC branch of our hybrid models. For example, the so-called W3 RGCs, for which the best stimulus found so far is a small dark moving spot [90], may not be "designed" to efficiently represent natural stimuli but rather to extract survival-relevant features (i.e., detecting aerial predators). Here, we could build models with different normative regularization or tasks (i.e., detecting predators in images of the sky) and would expect that this RGC type profits little from efficiently encoding natural statistics in the hybrid model. In this way, we may be able to discover the computational functions of specific cell types. Studying coding strategies across RGC types could contribute an important biological perspective to the perennial debate between efficient coding [91] and feature detection [67] proponents.

Normative network regularization as a framework for studying neural coding

In this study, we regularized the filters of a SI model with a normative EC model to predict visually-evoked responses of cells in the retina, which could be seen as a multitask learning model [92]. This approach is not limited to a combination of EC and SI, for example, Yamins et al. [11] used a model trained on an image categorization task for neural prediction in a sequential way. Some forms of normative regularization have also been discussed and/or applied in earlier work. For example, Deneve and Chalk [93] discussed the relations between SI (encoding) models and EC, and argued that the latter may promote shifting the focus in SI from the single-cell to the population level. The integration of stimulus-oriented approaches

(such as EC) for discriminative tasks (such as object recognition) was proposed by Turner et al. [15]. Later, Teti et al. [94] employed sparse coding with lateral inhibition in simulations of neuronal activation in visual cortex. More recently, Młynarski et al. [42] presented a probabilistic framework combining normative priors with statistical inference and demonstrated the usefulness of this approach for the analysis of diverse neuroscientific datasets. However, their work was rather conceptual, with the datasets they used being either simulated or low-dimensional. Notably, they tested their framework on pre-fit retinal RFs, but not directly on actual RGC stimulus-response data. Compared to their framework, our method does not require marginalization across all parameter space to estimate optimality and could be applied to more general or complex inference problems. Hence, our work not only provides further evidence to the feasibility of combining coding principles for identification of neural response properties on high-dimensional data, it also demonstrates the benefits of leveraging natural scene statistics for neural prediction. However, compared to the framework by Młynarski et al., with our approach it is more difficult to conduct rigorous statistical tests of normative theory.

We expect that our hybrid modeling strategy for prediction of responses to noise stimuli may also work for different processing stages along the early visual pathway (and potentially other modalities, e.g., sound). This said, however, one needs to keep in mind that different stages along the visual pathway have different tasks and constraints, and, thus, likely incorporate different efficient coding principles: For instance, the retinal hardware is space-limited and has to encode visual features in view of a bottleneck with limited bandwidth (optic nerve), whereas the primary visual cortex has comparably abundant resources which might serve for accurate probability estimation for behavioral tasks, such as novelty detection (discussed in [24, 95]). It is also worth to note that different visual processing stages (such as primary visual cortex vs. higher visual areas, or adaptation of visual coding to different behavioral states) may benefit from the hybrid modeling to a different degree, as efficient coding approaches learn filters that may be more relevant to stimulus-related features, but not high-level behavior goals (see discussion in [15]). Additionally, it would be interesting to compare our hybrid models with SI models regularized with other behavioral tasks such as object recognition (e.g., [11]) or predator detection (see above) for neural predictions along the ventral visual stream.

While this study focused on normative regularization for neural prediction task, it would be also interesting to infer EC principles from stimulus-response data. With our framework, a possible starting point could be to compare a normative criterion, such as image reconstruction fidelity, between the hybrid model and a stand-alone EC model. Such analysis could be extended by either evaluating the difference for coding principles with the use of the same stimulus-response data, or testing a normative criterion using different experimental datasets.

There is a long tradition of using SI models (reviewed in [3]) in predicting the responses of neurons to a great variety of stimuli (e.g., [2, 4, 18, 19, 96, 97]). Our results demonstrate how the EC hypothesis can be successfully leveraged as normative regularization for the identification of neural response properties when assessed through noise stimuli. Additionally, predicting the response to naturalistic stimuli may be more beneficial for learning biologically-plausible filters. More generally, the hybrid framework offers an opportunity to test different coding principles and unsupervised learning objectives with regards to experimental data for understanding neuronal processing.

Supporting information

S1 Fig. Training of 2D models. a. The noise stimulus (9 minutes in total) containing training and validation data (1 repeat) and test data (6 repeats). **b.** Model performance (mean)

based on validation data for SI-PCA and SI-DCT with different numbers of basis. SI-PCA and SI-DCT yielded best performance when using 16 and 4 bases, respectively (each model for n = 10 random seeds; error bars represent 2.5 and 97.5 percentiles with bootstrapping). **c.** Training loss as a function of training epochs for the hybrid model ($Input_{EC}$, natural scenes) with different weights (w), indicated by color (right). **d.** Model performance based on validation data (with linear correlation coefficient as metric) during the *hybrid-natural* model training with different weights (colors as in (c)). As weight decreased from 1 to 0.2, more training epochs were needed to reach the best performance. The hybrid model performed best for $w = 0.2$. Note that the hybrid model showed a slower change in correlation coefficient (CC) around the peak at $w = 0.2$ (compared to $w = 1$), demonstrating the regularization effects of the EC branch on the hybrid model. **e.** Scatter plots for model predictions based on test data at a particular seed (each dot representing one neuron). Hybrid with natural scenes as $input_{EC}$ ($w = 0.2$) vs. SI, SI with PCA basis (16 bases), SI with DCT basis (4 bases), *hybrid-pha-scr* ($w = 0.3$) and *hybrid-noise* ($w = 0.4$). **f.** Upper: Three representative GCL cell responses (gray traces) to noise stimulus together with predictions of the best performing models on test data (black, SI; blue, SI with PCA basis; cyan, SI with DCT basis; red, hybrid w/ natural scenes as input in EC path; brown, hybrid w/ phase-scrambled scenes as input in EC path; magenta, hybrid w/ noise as input in EC path). Lower: Learned spatio-temporal RFs of the example cells, visualized by SVD. Same random seed as in (e).
(TIF)

S2 Fig. Three-dimensional hybrid networks embedding natural movies. a,b. Illustration of SI network (a) with 3D spatio-temporal convolutional filter, and EC network (b), reconstructing the 7^{th} frame (at $t - 1$) based on 8 continuous frames ($t - 7$ to t; encoding the past, c). Combined as a hybrid network, the two branches were trained in parallel with shared 3D filters (all spatio-temporal filters were shared; $Input_{EC}$, 8-frame UV-green movie clip; $Output_{EC}$, reconstruction of the 7^{th} frame of $Input_{EC}$). **c.** Example for input/output of the EC model for encoding the past (left; also see b) and exemplary spatio-temporal convolutional filters when using natural movies as input to train the EC model alone (right). **d.** Example for input/output of the EC model for predicting the future, i.e., predicting the 8^{th} frame from the first 7 frames ($t - 7$ to $t - 1$) of the clip, and exemplary spatio-temporal filters when using natural movies as input to train the EC model alone. During preprocessing, the 8^{th} frame of input was set to the mean of the first 7 frames, for UV and green channel, respectively. Note that for stand-alone EC models, all temporal components of filters for past encoding were very similar while those for future prediction were much more diverse.
(TIF)

S3 Fig. Training of 3D hybrid models. a,b. Model performance (mean) based on validation data for hybrid models w/ natural movies as $input_{EC}$ (a), applying past encoding (*hybrid-natural-past*) or future prediction (*hybrid-natural-future*) and for different weights, and for the *SI-PCA* model (b) with different numbers of basis (each model for n = 10 random seeds). **c.** Scatter plots for model predictions based on test data at a particular seed (each dot representing one neuron). *hybrid-natural-past* ($w = 0.4$) vs. SI, *SI-PCA* (128 PCA bases) and *hybrid-natural-future* ($w = 0.4$). Error bars in (a)–(b) represent 2.5 and 97.5 percentiles with bootstrapping. Both 3D hybrid models performed similarly, with a peak in predictive performance on the validation data at around $w = 0.4$ (a). This value of w was higher than for the 2D hybrid models ($w = 0.2$; cf. Fig 3c). We also examined the low-pass filtering effects on the 3D SI model by using PCA filters (3D *SI-PCA*) and varying the number of basis (b). Like for the 2D case when varying the number of basis, we found a

maximum in performance on the validation data at 128 bases, which was larger than the 16 bases in the 2D case (cf. S1(b) Fig).
(TIF)

S4 Fig. Hybrid model for encoding neuronal responses to 30-Hz dense noise. To test hybrid models for different stimuli, we recorded neuronal responses to the 30-Hz dense noise in the ventral retina. We yielded n = 64 neurons after quality control (Methods), which were used to train the SI and hybrid networks. **a.** Model performance (mean) based on validation data for hybrid models (w/ natural movies as input_{EC}), applying encoding-past (*hybrid-natural-past*) or predicting-future (*hybrid-natural-future*) and for different weights. Each model for n = 10 random seeds. Both models with similar performance for all weights, peaking at $w = 0.7$. **b.** Model performance (mean) based on test data for SI, *hybrid-natural-past* ($w = 0.7$) and *hybrid-natural-future* ($w = 0.7$). Each model for n = 10 random seeds. The two hybrid models had better performance with smaller standard deviation compared the SI model ($p < 0.0001$ for SI and *hybrid-natural-past*, $p = 0.9992$ for *hybrid-natural-past* and *hybrid-natural-future*; two-sided permutation test, n = 10,000 repeats). **c.** R-squared (mean) of fitting a 2D Gaussian to all the spatial filters in UV stimulus channel (each model for n = 10 random seeds; $p < 0.0001$ for SI and *hybrid-natural-past*, $p = 0.9888$ for *hybrid-natural-past* and *hybrid-natural-future*; two-sided permutation test, n = 10,000 repeats). **d.** Learned spatio-temporal filters of the three representative cells, visualized by SVD. Note that because all neurons in this data set were recorded in the ventral retina, their responses were dominated by the UV channel. Different temporal filters in the UV channel were observed for these neurons (cf. the very similar temporal filters in the green channel for neurons' responses to 5-Hz noise in Figs 3b and 5a lower). **e.** Exemplary shared spatial and temporal filters of 3D models, visualized by SVD and for one random seed. Temporal: UV and green channels indicated by purple and green lines, respectively. Error bars in (a)–(c) represent 2.5 and 97.5 percentiles with bootstrapping.
(TIF)

S5 Fig. Confusion matrix for a trained random forest classifier. Normalized confusion matrix (true cell types against predicted cell types) for a trained random forest classifier evaluated on a test dataset (for details, see Methods). Dotted line indicates separation of 6 broad functional cell groups [47].
(TIF)

S6 Fig. Hybrid model for different cell types. a. Performance difference (mean) between *hybrid-natural-future* and SI based on test data for different cell types (each model for n = 10 random seeds). **b.** Performance difference (mean) between *hybrid-natural-future* and SI based on test data for DS and non-DS cells (each model for n = 10 random seeds). **c.** Cumulative histogram of model prediction difference between *hybrid-natural-future* ($w = 0.7$) and SI on test data, for DS (red) and non-DS cells, at one particular seed. **d.** Scatter plots for model predictions based on test data at a particular seed (each dot representing one neuron) for DS and non-DS cells and *hybrid-natural-past* ($w = 0.7$) vs. *hybrid-natural-future* ($w = 0.7$). Note that the predictions of two hybrid models were similar for most of neurons. **e.** Quality index (mean) for DS and non-DS cells based on responses to the repeated test sequences in the noise stimuli ($p = 0.2881$, two-sided permutation test, n = 10,000 repeats; for details, see Methods). **f.** Like (e) but for chirp responses ($p = 0.6714$, two-sided permutation test, n = 10,000 repeats). **g.** Like (e) but for bar stimulus responses ($p < 0.0001$, two-sided permutation test, n = 10,000 repeats). Error bars in (a),(b),(e)-(g) represent 2.5 and 97.5 percentiles with bootstrapping.
(TIF)

Acknowledgments

We thank Matthew Chalk, Dylan Paiton and Katrin Franke for helpful discussions, and Merle Harrer for excellent technical assistance.

Author Contributions

Conceptualization: Yongrong Qiu.

Formal analysis: Yongrong Qiu, David A. Klindt, Laura Busse, Matthias Bethge, Thomas Euler.

Funding acquisition: Laura Busse, Matthias Bethge, Thomas Euler.

Investigation: Yongrong Qiu, David A. Klindt, Klaudia P. Szatko, Laura Busse, Matthias Bethge, Thomas Euler.

Methodology: Yongrong Qiu, David A. Klindt, Klaudia P. Szatko, Dominic Gonschorek, Larissa Hoefling, Laura Busse, Matthias Bethge, Thomas Euler.

Project administration: Thomas Euler.

Resources: Timm Schubert, Thomas Euler.

Software: Yongrong Qiu, Dominic Gonschorek, Larissa Hoefling.

Supervision: Laura Busse, Matthias Bethge, Thomas Euler.

Validation: Yongrong Qiu, David A. Klindt.

Visualization: Yongrong Qiu, Dominic Gonschorek.

Writing – original draft: Yongrong Qiu, David A. Klindt, Laura Busse, Thomas Euler.

Writing – review & editing: Yongrong Qiu, David A. Klindt, Klaudia P. Szatko, Dominic Gonschorek, Larissa Hoefling, Timm Schubert, Laura Busse, Matthias Bethge, Thomas Euler.

References

1. Stevenson IH, Kording KP. How advances in neural recording affect data analysis. Nature neuroscience. 2011; 14(2):139–142. https://doi.org/10.1038/nn.2731 PMID: 21270781
2. Chichilnisky E. A simple white noise analysis of neuronal light responses. Network: Computation in Neural Systems. 2001; 12(2):199–213. PMID: 11405422
3. Wu MCK, David SV, Gallant JL. Complete functional characterization of sensory neurons by system identification. Annu Rev Neurosci. 2006; 29:477–505. https://doi.org/10.1146/annurev.neuro.29.051605.113024 PMID: 16776594
4. Pillow JW, Shlens J, Paninski L, Sher A, Litke AM, Chichilnisky E, et al. Spatio-temporal correlations and visual signalling in a complete neuronal population. Nature. 2008; 454(7207):995–999. https://doi.org/10.1038/nature07140 PMID: 18650810
5. Marmarelis V. Analysis of physiological systems: The white-noise approach. Springer Science & Business Media; 2012.
6. Koelling ME, Nykamp DQ. Computing linear approximations to nonlinear neuronal response. Network: Computation in Neural Systems. 2008; 19(4):286–313. https://doi.org/10.1080/09548980802503139 PMID: 18991145
7. Gollisch T, Meister M. Eye smarter than scientists believed: neural computations in circuits of the retina. Neuron. 2010; 65(2):150–164. https://doi.org/10.1016/j.neuron.2009.12.009 PMID: 20152123
8. Real E, Asari H, Gollisch T, Meister M. Neural circuit inference from function to structure. Current Biology. 2017; 27(2):189–198. https://doi.org/10.1016/j.cub.2016.11.040 PMID: 28065610

9. Willmore B, Prenger RJ, Wu MCK, Gallant JL. The berkeley wavelet transform: a biologically inspired orthogonal wavelet transform. Neural computation. 2008; 20(6):1537–1564. https://doi.org/10.1162/neco.2007.05-07-513 PMID: 18194102
10. Maheswaranathan N, Kastner DB, Baccus SA, Ganguli S. Inferring hidden structure in multilayered neural circuits. PLoS computational biology. 2018; 14(8):e1006291. https://doi.org/10.1371/journal.pcbi.1006291 PMID: 30138312
11. Yamins DL, Hong H, Cadieu CF, Solomon EA, Seibert D, DiCarlo JJ. Performance-optimized hierarchical models predict neural responses in higher visual cortex. Proceedings of the National Academy of Sciences. 2014; 111(23):8619–8624. https://doi.org/10.1073/pnas.1403112111 PMID: 24812127
12. Güçlü U, van Gerven MA. Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. Journal of Neuroscience. 2015; 35(27):10005–10014. https://doi.org/10.1523/JNEUROSCI.5023-14.2015 PMID: 26157000
13. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015; 521(7553):436–444. https://doi.org/10.1038/nature14539 PMID: 26017442
14. Hassabis D, Kumaran D, Summerfield C, Botvinick M. Neuroscience-inspired artificial intelligence. Neuron. 2017; 95(2):245–258. https://doi.org/10.1016/j.neuron.2017.06.011 PMID: 28728020
15. Turner MH, Giraldo LGS, Schwartz O, Rieke F. Stimulus-and goal-oriented frameworks for understanding natural vision. Nature neuroscience. 2019; 22(1):15–24. https://doi.org/10.1038/s41593-018-0284-0 PMID: 30531846
16. Richards BA, Lillicrap TP, Beaudoin P, Bengio Y, Bogacz R, Christensen A, et al. A deep learning framework for neuroscience. Nature neuroscience. 2019; 22(11):1761–1770. https://doi.org/10.1038/s41593-019-0520-2 PMID: 31659335
17. Yamins DL, DiCarlo JJ. Using goal-driven deep learning models to understand sensory cortex. Nature neuroscience. 2016; 19(3):356–365. https://doi.org/10.1038/nn.4244 PMID: 26906502
18. McIntosh L, Maheswaranathan N, Nayebi A, Ganguli S, Baccus S. Deep learning models of the retinal response to natural scenes. Advances in neural information processing systems. 2016; 29:1369–1377. PMID: 28729779
19. Klindt D, Ecker AS, Euler T, Bethge M. Neural system identification for large populations separating "what" and "where". Advances in Neural Information Processing Systems. 2017; 31:3506–3516.
20. Bashivan P, Kar K, DiCarlo JJ. Neural population control via deep image synthesis. Science. 2019; 364 (6439). https://doi.org/10.1126/science.aav9436 PMID: 31048462
21. Ponce CR, Xiao W, Schade PF, Hartmann TS, Kreiman G, Livingstone MS. Evolving images for visual neurons using a deep generative network reveals coding principles and neuronal preferences. Cell. 2019; 177(4):999–1009. https://doi.org/10.1016/j.cell.2019.04.005 PMID: 31051108
22. Walker EY, Sinz FH, Cobos E, Muhammad T, Froudarakis E, Fahey PG, et al. Inception loops discover what excites neurons most using deep predictive models. Nature neuroscience. 2019; 22(12):2060–2065. https://doi.org/10.1038/s41593-019-0517-x PMID: 31686023
23. Baden T, Euler T, Berens P. Understanding the retinal basis of vision across species. Nature Reviews Neuroscience. 2020; 21(1):5–20. https://doi.org/10.1038/s41583-019-0242-1 PMID: 31780820
24. Barlow HB, et al. Possible principles underlying the transformation of sensory messages. Sensory communication. 1961; 1(01).
25. Simoncelli EP, Olshausen BA. Natural image statistics and neural representation. Annual review of neuroscience. 2001; 24(1):1193–1216. https://doi.org/10.1146/annurev.neuro.24.1.1193 PMID: 11520932
26. Switkes E, Mayer MJ, Sloan JA. Spatial frequency analysis of the visual environment: Anisotropy and the carpentered environment hypothesis. Vision research. 1978; 18(10):1393–1399. https://doi.org/10.1016/0042-6989(78)90232-8 PMID: 726283
27. Xu X, Collins CE, Khaytin I, Kaas JH, Casagrande VA. Unequal representation of cardinal vs. oblique orientations in the middle temporal visual area. Proceedings of the National Academy of Sciences. 2006; 103(46):17490–17495. https://doi.org/10.1073/pnas.0608502103 PMID: 17088527
28. Girshick AR, Landy MS, Simoncelli EP. Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. Nature neuroscience. 2011; 14(7):926–932. https://doi.org/10.1038/nn.2831 PMID: 21642976
29. Laughlin S. A simple coding procedure enhances a neuron's information capacity. Zeitschrift für Naturforschung c. 1981; 36(9–10):910–912. https://doi.org/10.1515/znc-1981-9-1040 PMID: 7303823
30. van Hateren JH, Ruderman DL. Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. Proceedings of the Royal Society

of London Series B: Biological Sciences. 1998; 265(1412):2315–2320. https://doi.org/10.1098/rspb.1998.0577 PMID: 9881476

31. Roy S, Jun NY, Davis EL, Pearson J, Field GD. Inter-mosaic coordination of retinal receptive fields. Nature. 2021; 592(7854):409–413. https://doi.org/10.1038/s41586-021-03317-5 PMID: 33692544
32. Atick JJ, Redlich AN. Towards a theory of early visual processing. Neural computation. 1990; 2(3):308–320. https://doi.org/10.1162/neco.1990.2.3.308
33. Atick JJ. Could information theory provide an ecological theory of sensory processing? Network: Computation in neural systems. 1992; 3(2):213–251. https://doi.org/10.1088/0954-898X_3_2_009
34. Li Z. Understanding vision: theory, models, and data. Oxford University Press, USA; 2014.
35. Enroth-Cugell C, Robson JG. The contrast sensitivity of retinal ganglion cells of the cat. The Journal of physiology. 1966; 187(3):517–552. https://doi.org/10.1113/jphysiol.1966.sp008107 PMID: 16783910
36. Ballard DH. Modular learning in neural networks. In: AAAI; 1987. p. 279–284.
37. Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science. 2006; 313(5786):504–507. https://doi.org/10.1126/science.1127647 PMID: 16873662
38. Ocko S, Lindsey J, Ganguli S, Deny S. The emergence of multiple retinal cell types through efficient coding of natural movies. Advances in Neural Information Processing Systems. 2018; 32:9389–9400.
39. Qiu Y, Zhao Z, Klindt D, Kautzky M, Szatko KP, Schaeffel F, et al. Natural environment statistics in the upper and lower visual field are reflected in mouse retinal specializations. Current Biology. 2021;. https://doi.org/10.1016/j.cub.2021.05.017 PMID: 34107304
40. Paiton DM, Frye CG, Lundquist SY, Bowen JD, Zarcone R, Olshausen BA. Selectivity and robustness of sparse coding networks. Journal of Vision. 2020; 20(12):10–10. https://doi.org/10.1167/jov.20.12.10 PMID: 33237290
41. Eichhorn J, Sinz F, Bethge M. Natural image coding in V1: how much use is orientation selectivity? PLoS computational biology. 2009; 5(4):e1000336. https://doi.org/10.1371/journal.pcbi.1000336 PMID: 19343216
42. Młynarski W, Hledík M, Sokolowski TR, Tkačik G. Statistical analysis and optimality of neural systems. Neuron. 2021; 109(7):1227–1241. https://doi.org/10.1016/j.neuron.2021.01.020 PMID: 33592180
43. Briggman KL, Euler T. Bulk electroporation and population calcium imaging in the adult mammalian retina. Journal of neurophysiology. 2011; 105(5):2601–2609. https://doi.org/10.1152/jn.00722.2010 PMID: 21346205
44. Szatko KP, Korympidou MM, Ran Y, Berens P, Dalkara D, Schubert T, et al. Neural circuits in the mouse retina support color vision in the upper visual field. Nature communications. 2020; 11(1):1–14. https://doi.org/10.1038/s41467-020-17113-8 PMID: 32661226
45. Euler T, Hausselt SE, Margolis DJ, Breuninger T, Castell X, Detwiler PB, et al. Eyecup scope—optical recordings of light stimulus-evoked fluorescence signals in the retina. Pflügers Archiv-European Journal of Physiology. 2009; 457(6):1393–1414. https://doi.org/10.1007/s00424-008-0603-5 PMID: 19023590
46. Euler T, Franke K, Baden T. Studying a light sensor with light: multiphoton imaging in the retina. In: Multiphoton Microscopy. Springer; 2019. p. 225–250.
47. Baden T, Berens P, Franke K, Rosón MR, Bethge M, Euler T. The functional diversity of retinal ganglion cells in the mouse. Nature. 2016; 529(7586):345–350. https://doi.org/10.1038/nature16468 PMID: 26735013
48. Franke K, Chagas AM, Zhao Z, Zimmermann MJ, Bartel P, Qiu Y, et al. An arbitrary-spectrum spatial visual stimulator for vision research. elife. 2019; 8:e48779. https://doi.org/10.7554/eLife.48779 PMID: 31545172
49. Breiman L. Random forests. Machine learning. 2001; 45(1):5–32. https://doi.org/10.1023/A:1010933404324
50. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 2011; 12:2825–2830.
51. Sun L, Jia K, Yeung DY, Shi BE. Human action recognition using factorized spatio-temporal convolutional networks. In: Proceedings of the IEEE international conference on computer vision; 2015. p. 4597–4605.
52. Tran D, Wang H, Torresani L, Ray J, LeCun Y, Paluri M. A closer look at spatiotemporal convolutions for action recognition. In: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition; 2018. p. 6450–6459.
53. Vincent BT, Baddeley RJ. Synaptic energy efficiency in retinal processing. Vision research. 2003; 43(11):1285–1292. https://doi.org/10.1016/S0042-6989(03)00096-8 PMID: 12726834
54. Doi E, Lewicki MS. A theory of retinal population coding. Advances in neural information processing systems. 2007; 19:353.

55. Van Rossum M, O'Brien BJ, Smith RG. Effects of noise on the spike timing precision of retinal ganglion cells. Journal of neurophysiology. 2003; 89(5):2406–2419. https://doi.org/10.1152/jn.01106.2002 PMID: 12740401

56. Field DJ. What is the goal of sensory coding? Neural computation. 1994; 6(4):559–601. https://doi.org/10.1162/neco.1994.6.4.559

57. Hubel DH, Wiesel TN. Receptive fields of single neurones in the cat's striate cortex. The Journal of physiology. 1959; 148(3):574–591. https://doi.org/10.1113/jphysiol.1959.sp006308 PMID: 14403679

58. Marr D, Hildreth E. Theory of edge detection. Proceedings of the Royal Society of London Series B Biological Sciences. 1980; 207(1167):187–217. PMID: 6102765

59. Schlamp CL, Montgomery AD, Mac Nair CE, Schuart C, Willmer DJ, Nickells RW. Evaluation of the percentage of ganglion cells in the ganglion cell layer of the rodent retina. Molecular vision. 2013; 19:1387. PMID: 23825918

60. Jacobs GH, Williams GA, Fenwick JA. Influence of cone pigment coexpression on spectral sensitivity and color vision in the mouse. Vision research. 2004; 44(14):1615–1622. https://doi.org/10.1016/j.visres.2004.01.016 PMID: 15135998

61. Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. In: European conference on computer vision. Springer; 2014. p. 818–833.

62. Franke K, Berens P, Schubert T, Bethge M, Euler T, Baden T. Inhibition decorrelates visual feature representations in the inner retina. Nature. 2017; 542(7642):439–444. https://doi.org/10.1038/nature21394 PMID: 28178238

63. Soodak RE. Two-dimensional modeling of visual receptive fields using Gaussian subunits. Proceedings of the National Academy of Sciences. 1986; 83(23):9259–9263. https://doi.org/10.1073/pnas.83.23.9259 PMID: 3466186

64. Palmer SE, Marre O, Berry MJ, Bialek W. Predictive information in a sensory population. Proceedings of the National Academy of Sciences. 2015; 112(22):6908–6913. https://doi.org/10.1073/pnas.1506855112

65. Bialek W, Van Steveninck RRDR, Tishby N. Efficient representation as a design principle for neural coding and computation. In: 2006 IEEE international symposium on information theory. IEEE; 2006. p. 659–663.

66. Chalk M, Marre O, Tkačik G. Toward a unified theory of efficient, predictive, and sparse coding. Proceedings of the National Academy of Sciences. 2018; 115(1):186–191. https://doi.org/10.1073/pnas.1711114115

67. Lettvin JY, Maturana HR, McCulloch WS, Pitts WH. What the frog's eye tells the frog's brain. Proceedings of the IRE. 1959; 47(11):1940–1951. https://doi.org/10.1109/JRPROC.1959.287207

68. Bae JA, Mu S, Kim JS, Turner NL, Tartavull I, Kemnitz N, et al. Digital museum of retinal ganglion cells with dense anatomy and physiology. Cell. 2018; 173(5):1293–1306. https://doi.org/10.1016/j.cell.2018.04.040 PMID: 29775596

69. Tran NM, Shekhar K, Whitney IE, Jacobi A, Benhar I, Hong G, et al. Single-cell profiles of retinal ganglion cells differing in resilience to injury reveal neuroprotective genes. Neuron. 2019; 104(6):1039–1055. https://doi.org/10.1016/j.neuron.2019.11.006 PMID: 31784286

70. Goetz J, Jessen ZF, Jacobi A, Mani A, Cooler S, Greer D, et al. Unified classification of mouse retinal ganglion cells using function, morphology, and gene expression. Morphology, and Gene Expression. 2021;.

71. Barlow HB, Hill RM. Selective sensitivity to direction of movement in ganglion cells of the rabbit retina. Science. 1963; 139(3553):412–412. https://doi.org/10.1126/science.139.3553.412 PMID: 13966712

72. Olshausen BA, Field DJ. How close are we to understanding V1? Neural computation. 2005; 17(8):1665–1699. https://doi.org/10.1162/0899766054026639 PMID: 15969914

73. Heitman A, Brackbill N, Greschner M, Sher A, Litke AM, Chichilnisky E. Testing pseudo-linear models of responses to natural scenes in primate retina. bioRxiv. 2016; p. 045336.

74. Rust NC, Movshon JA. In praise of artifice. Nature neuroscience. 2005; 8(12):1647–1650. https://doi.org/10.1038/nn1606 PMID: 16306892

75. Touryan J, Felsen G, Dan Y. Spatial structure of complex cell receptive fields measured with natural images. Neuron. 2005; 45(5):781–791. https://doi.org/10.1016/j.neuron.2005.01.029 PMID: 15748852

76. Jacoby J, Schwartz GW. Three small-receptive-field ganglion cells in the mouse retina are distinctly tuned to size, speed, and object motion. Journal of Neuroscience. 2017; 37(3):610–625. https://doi.org/10.1523/JNEUROSCI.2804-16.2016 PMID: 28100743

77. Rousso DL, Qiao M, Kagan RD, Yamagata M, Palmiter RD, Sanes JR. Two pairs of ON and OFF retinal ganglion cells are defined by intersectional patterns of transcription factor expression. Cell reports. 2016; 15(9):1930–1944. https://doi.org/10.1016/j.celrep.2016.04.069 PMID: 27210758

78. Bleckert A, Schwartz GW, Turner MH, Rieke F, Wong RO. Visual space is represented by nonmatching topographies of distinct mouse retinal ganglion cell types. Current Biology. 2014; 24(3):310–315. https://doi.org/10.1016/j.cub.2013.12.020 PMID: 24440397

79. Brown SP, He S, Masland RH. Receptive field microstructure and dendritic geometry of retinal ganglion cells. Neuron. 2000; 27(2):371–383. https://doi.org/10.1016/S0896-6273(00)00044-1 PMID: 10985356

80. Hénaff OJ, Goris RL, Simoncelli EP. Perceptual straightening of natural videos. Nature neuroscience. 2019; 22(6):984–991. https://doi.org/10.1038/s41593-019-0377-4 PMID: 31036946

81. Hénaff OJ, Bai Y, Charlton JA, Nauhaus I, Simoncelli EP, Goris RL. Primary visual cortex straightens natural video trajectories. Nature communications. 2021; 12(1):5982. https://doi.org/10.1038/s41467-021-25939-z PMID: 34645787

82. Saxe A, Nelli S, Summerfield C. If deep learning is the answer, what is the question? Nature Reviews Neuroscience. 2021; 22(1):55–67. https://doi.org/10.1038/s41583-020-00395-8 PMID: 33199854

83. Rao RP, Ballard DH. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nature neuroscience. 1999; 2(1):79–87. https://doi.org/10.1038/4580 PMID: 10195184

84. Hosoya T, Baccus SA, Meister M. Dynamic predictive coding by the retina. Nature. 2005; 436 (7047):71–77. https://doi.org/10.1038/nature03689 PMID: 16001064

85. Johnston J, Seibel SH, Darnet LSA, Renninger S, Orger M, Lagnado L. A retinal circuit generating a dynamic predictive code for oriented features. Neuron. 2019; 102(6):1211–1222. https://doi.org/10.1016/j.neuron.2019.04.002 PMID: 31054873

86. Sinz FH, Ecker AS, Fahey PG, Walker EY, Cobos E, Froudarakis E, et al. Stimulus domain transfer in recurrent models for large scale cortical population prediction on video. BioRxiv. 2018; p. 452672.

87. Olshausen BA, Field DJ. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature. 1996; 381(6583):607–609. https://doi.org/10.1038/381607a0 PMID: 8637596

88. Karklin Y, Simoncelli EP. Efficient coding of natural images with a population of noisy linear-nonlinear neurons. Advances in neural information processing systems. 2011; 24:999. PMID: 26273180

89. Doi E, Lewicki MS. A simple model of optimal population coding for sensory systems. PLoS computational biology. 2014; 10(8):e1003761. https://doi.org/10.1371/journal.pcbi.1003761 PMID: 25121492

90. Zhang Y, Kim IJ, Sanes JR, Meister M. The most numerous ganglion cell type of the mouse retina is a selective feature detector. Proceedings of the National Academy of Sciences. 2012; 109(36):E2391–E2398. https://doi.org/10.1073/pnas.1211547109 PMID: 22891316

91. Barlow HB. Summation and inhibition in the frog's retina. The Journal of physiology. 1953; 119(1):69–88. https://doi.org/10.1113/jphysiol.1953.sp004829 PMID: 13035718

92. Caruana R. Multitask learning. Machine learning. 1997; 28(1):41–75. https://doi.org/10.1023/A:1007379606734

93. Deneve S, Chalk M. Efficiency turns the table on neural encoding, decoding and noise. Current Opinion in Neurobiology. 2016; 37:141–148. https://doi.org/10.1016/j.conb.2016.03.002 PMID: 27065340

94. Teti M, Meyer E, Kenyon G. Can Lateral Inhibition for Sparse Coding Help Explain V1 Neuronal Responses To Natural Stimuli? In: 2020 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI). IEEE; 2020. p. 120–124.

95. Barlow H. Redundancy reduction revisited. Network: computation in neural systems. 2001; 12(3):241. https://doi.org/10.1080/net.12.3.241.253 PMID: 11563528

96. Vintch B, Movshon JA, Simoncelli EP. A convolutional subunit model for neuronal responses in macaque V1. Journal of Neuroscience. 2015; 35(44):14829–14841. https://doi.org/10.1523/JNEUROSCI.2815-13.2015 PMID: 26538653

97. Cadena SA, Denfield GH, Walker EY, Gatys LA, Tolias AS, Bethge M, et al. Deep convolutional models improve predictions of macaque V1 responses to natural images. PLoS computational biology. 2019; 15(4):e1006897. https://doi.org/10.1371/journal.pcbi.1006897 PMID: 31013278

Study II: Removing Inter-Experimental Variability from Functional Data in Systems Neuroscience:

Published as "Gonschorek et al. 'Removing Inter-experimental Variability from Functional Data in Systems Neuroscience'; Advances in Neural Information Processing Systems (NeurIPS 2021), 34. (2021)."

Removing Inter-Experimental Variability from Functional Data in Systems Neuroscience

Dominic Gonschorek[*1]
dominic.gonschorek@cin.uni-tuebingen.de

Larissa Höfling[*,1]
larissa.hoefling@uni-tuebingen.de

Klaudia P. Szatko[1]
klaudia.szatko@tuebingen.mpg.de

Katrin Franke[1]
katrin.franke@cin.uni-tuebingen.de

Timm Schubert[1]
timm.schubert@cin.uni-tuebingen.de

Benjamin A. Dunn[2]
benjamin.dunn@ntnu.no

Philipp Berens[1]
philipp.berens@uni-tuebingen.de

David A. Klindt[‡,2]
klindt.david@gmail.com

Thomas Euler[‡,1]
thomas.euler@cin.uni-tuebingen.de

Abstract

Integrating data from multiple experiments is common practice in systems neuroscience but it requires *inter-experimental variability* to be negligible compared to the biological signal of interest. This requirement is rarely fulfilled; systematic changes between experiments can drastically affect the outcome of complex analysis pipelines. Modern machine learning approaches designed to adapt models across multiple data domains offer flexible ways of removing inter-experimental variability where classical statistical methods often fail. While applications of these methods have been mostly limited to single-cell genomics, in this work, we develop a theoretical framework for domain adaptation in systems neuroscience. We implement this in an adversarial optimization scheme that removes inter-experimental variability while preserving the biological signal. We compare our method to previous approaches on a large-scale dataset of two-photon imaging recordings of retinal bipolar cell responses to visual stimuli. This dataset provides a unique benchmark as it contains biological signal from well-defined cell types that is obscured by large inter-experimental variability. In a supervised setting, we compare the generalization performance of cell type classifiers across experiments, which we validate with anatomical cell type distributions from electron microscopy data. In an unsupervised setting, we remove inter-experimental variability from data which can then be fed into arbitrary downstream analyses. In both settings, we find that our method achieves the best trade-off between removing inter-experimental variability and preserving biological signal. Thus, we offer a flexible approach to remove inter-experimental variability and integrate datasets across experiments in systems neuroscience. Code available at `https://github.com/eulerlab/rave`.

[*‡]Equal contributions, [1] University of Tübingen, [2] Norwegian University of Science and Technology.

35th Conference on Neural Information Processing Systems (NeurIPS 2021).

1 Introduction

Systems neuroscientists are often concerned with identifying and characterizing how properties of neurons vary along certain dimensions of interest. Differences in these properties between neurons form the basis for sorting them into discrete categories. Both the advance of large-scale data acquisition techniques in experimental neuroscience as well as the development of more efficient and powerful data analysis methods allow collecting and analyzing datasets of increasing size; and hence the discovery of more subtle variations in neural function between cell types [e.g. 1–4]. However, as data acquisition is often an incremental process, it has become common practice to pool data from multiple experiments. This practice ignores variability in the data stemming from external factors, which include non-biological ones (e.g. sample handling resulting in small differences in tissue quality, or temperature fluctuations affecting the rates of biochemical processes) but potentially also unforeseen biological ones (e.g. subtle genetic variations) [3, 5, 6]. Such variability due to external factors, here referred to as *inter-experimental variability*, can confound and obscure the biological signals of interest. In some cases, the source of inter-experimental variability is known and can be modeled [5], but if this is not the case, a method for removing it from the data is required.

The issue of inter-experimental variability in systems neuroscience is analogous to the problem of *domain shift* in machine learning, where the data distribution changes between training ('source') and test ('target') data, causing an algorithm to fail when deployed on data from an unseen target domain [7–10]. Methods that address this issue have to perform some form of *domain adaptation*, i.e. adapting the algorithm to work both on the training as well as some (usually unseen) test domain [11]. In single-cell genomics, a number of different studies have proposed methods for removing inter-experimental variability (see Section 2), but related works in systems neuroscience are lacking, despite the recognized need for such approaches [3, 5]. Here, we contribute to closing this gap as follows:

- We cast the removal of inter-experimental variability from functional data in systems neuroscience in the theoretical framework of domain adaptation (Figure 1 and Section 3).
- We adapt and evaluate different approaches and demonstrate improved performance of cell type assignment, while preserving the biological signals of interest (Table 1 and Figure 4).
- We demonstrate that our method produces cell type predictions on a new dataset that are best aligned with anatomical data (Table 2 and Figure 5).
- Finally, we showcase in a downstream analysis that the corrected data (Figure 3) clearly exhibits biological effects that were obscured by inter-experimental variability (Figure 6).

2 Related Work

As mentioned before, few studies have proposed specialized solutions to the issue of inter-experimental variability in systems neuroscience. Two studies have approached the problem of temporal alignment of neural responses across experiments. Zhao et al. [5] proposed a solution to deal with the specific effects of temperature fluctuations on the response kinetics of retinal neurons by modeling them explicitly. Williams et al. [12] proposed a more general method for the temporal alignment of data across trials or recording sessions. Other studies have suggested models of neural function that integrate data across experiments. Shah et al. [3] build encoding models to predict the responses of retinal ganglion cells across different experiments [see also 13] and compare it to covariates such as the gender of an animal. Sorochynsky et al. [14] propose a way to measure noise correlations in each recording and integrate those into models of neural populations of a specific cell type. This latter approach is complementary to our method because it allows the study of the structure and function [see also 15] of noise correlations, which we discard as nuisance variability. Somewhat related to the example application in our paper, Jouty et al. [16] suggested a method to perform non-parametric physiological classification of retinal ganglion cells in the mouse retina while trying to find matching clusters of cell types across experiments. Crucially, all of these approaches offer specialized solutions that do not represent general purpose correction methods.

In single-cell genomics, a number of approaches for removing inter-experimental variability from data have been developed [17–24]. Two such methods are *Harmony* [25] and *scGen* [26]. *Harmony* performs iterative clustering using a variant of soft k-means until convergence to align cells from different datasets in a joint embedding. *scGen*, on the other hand, combines a variational autoencoder

adapted for scRNA-seq data with latent space arithmetics to predict gene expression, while removing inter-experimental variability between datasets. In this paper, we compare our approach against these methods as they have been found to perform particularly well in two benchmarking studies [27, 28].

3 Theoretical Framework

The generative process of data denoted by a random variable X with image $\mathcal{X}$ is depicted in Figure 1. The biological signal shared across experiments (e.g. variation due to cell types) is represented by a random variable S ('signal') with image $\mathcal{S}$. We define D ('domain') as a random variable with image $\mathcal{D}$ that represents inter-experimental variability. Now, our objective is to learn a function f that transforms the data into a new random variable $Z := f(X)$ with image $\mathcal{Z}$. Importantly, we distinguish two settings: (*i*) *unsupervised* — where S is unknown and we simply try to retain in Z as much information about the data as possible while removing inter-experimental variability; (*ii*) *supervised* — where S is known and we additionally try to retain in Z as much information about S as possible. These objectives can be formulated in terms of mutual information, giving the unsupervised loss function

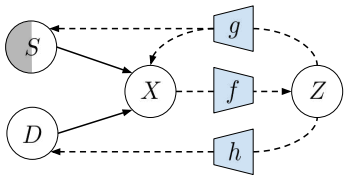

Figure 1: **Problem Setting**. Arrows represent given (solid lines) and modeled (dashed lines) relations. Capital letters denote random variables, small letters transformations (see Section 3). The setting with known S (white circle) is supervised, while unknown S (gray circle) is an unsupervised setting.

$$\mathcal{L} = I(Z; D) - I(Z; X) \tag{1}$$

and, provided knowledge about S, we obtain the supervised loss function

$$\mathcal{L}_+ = \mathcal{L} - I(Z; S). \tag{2}$$

Now, $I(Z; D)$ attains its minimum for $I(f(X); D) = 0$ because of the non-negativity of mutual information. And $I(Z; S)$ attains its maximum for $I(f(X); S) = I(X; S)$ because of the data processing inequality. If f were a bijection, it would follow that $I(f(X); S) = I(X; S)$, but also $I(f(X); D) = I(X; D)$. But by assumption, $I(X; D) > 0$ (otherwise there is no inter-experimental variability and we are done) and so we would have $I(f(X); D) > 0$. Thus, at the minimum of $I(Z; D)$, f cannot be a bijection. Generally, if there is an interaction between recording, signal and domain i.e. $I(X; S; D) \neq 0$, then there will be a *trade-off* between maximizing $I(f(X); S)$ and minimizing $I(f(X); D)$. This trade-off becomes even more apparent in the unsupervised setting where $I(f(X); X)$ and $I(f(X); D)$ are clearly competing.

Mutual information quantifies the dependence between two variables but it is difficult to estimate [29–32]. Instead, we measure dependency through nonlinear regression with an appropriate distance metric d.[2] Usually D is a discrete random variable indicating the experiment of a recording, and so, to estimate $I(Z; D)$, we can perform classification with a classifier function $h : \mathcal{Z} \to \mathcal{D}$, minimizing the standard cross-entropy $d_{\text{CE}}(h(Z), D)$ (Figure 1). Lemma 10 in [34] shows that this gives a variational lower bound to $I(Z; D)$ [see also 35]. In some cases S may also be discrete (e.g. cell types) and we can do the same, in other cases it might be a (high-dimensional) continuous random variable and so, to approximate $I(Z; S)$, we can perform regression, minimizing the mean squared error $d_{\text{MSE}}(g(Z), S)$. Similarly, in the unsupervised setting, to approximate $I(Z; X)$, we minimize $d_{\text{MSE}}(g(Z), X)$. To keep notation simple, in the unsupervised setting we define the mapping $g : \mathcal{Z} \to \mathcal{X}$, and in the supervised setting $g : \mathcal{Z} \to \mathcal{X} \times \mathcal{S}$. Putting this together, in the unsupervised setting our objective is

$$\min \mathcal{L} \longrightarrow \min_{h} \max_{g,f} \lambda\, d(h(f(X)), D) - d(g(f(X)), X) \tag{3}$$

where we have introduced a hyperparameter λ that mitigates the trade-off discussed above. In the supervised setting our objective becomes

$$\min \mathcal{L}_+ \longrightarrow \min_{h} \max_{g,f} \lambda\, d(h(f(X)), D) - d(g(f(X)), (X, S)). \tag{4}$$

[2] If the joint probability density of two random variables is a bivariate normal distribution, then the mutual information is proportional to their linear correlation [33].

In both equations we find a $\min - \max$ optimization, where h is trying to predict D from Z, tightening the lower bound on $I(Z; D)$ [see 34], while f is trying to prevent that by removing information about D from Z, effectively lowering $I(Z; D)$. Practically, this optimization scheme has become a standard adversarial setting in machine learning, for instance, in the training of generative adversarial networks [36] or for discriminative adversarial domain adaptation [37].

4 Methods

4.1 Datasets

To test our model approach, we use two datasets of two-photon imaging recordings [38–40] from the 14 mouse retinal bipolar cell (BC) [41] types' responses to two visual stimuli, a local and full-field chirp stimulus (Figure 3). The axon terminals of BCs stratify at distinct, cell type-specific depths within the second synaptic layer of the retina, the inner plexiform layer (IPL). The functional BC data were obtained by imaging the glutamate output at their axon terminals using the genetically encoded glutamate-sensing fluorescent reporter iGluSnFR [42]. In our study, we refer to these two datasets as *A* [2] and *B* [5] (for further preprocessing see Appendix).

In [2], an anatomy-guided functional clustering approach to group the BCs into the 14 functional types was applied to dataset *A*, thus providing functional reference cell type labels, which do not exist for dataset *B*. However, even if both datasets recorded the same cell types, they suffer from inter-experimental variability making it difficult to match and, for example, to use dataset *A* to predict the cell type labels for dataset *B*. We discuss potential sources of inter-experimental variability in the Appendix. For preprocessing, both BC datasets are high-pass filtered above 0.1 Hz (to remove the trends of decreasing fluorescence signal over time) and resampled to 30 Hz. Each cell's response is normalized to zero mean and standard deviation one. In addition, to ensure high quality responses, only cells with a sufficient response quality are used (for details about quality criterion see [2]).

4.2 Models

All methods transform the data, either into a low-dimensional embedding $z \in \mathcal{Z}$ or directly into a reconstruction $\hat{x} \in \mathcal{X}$ from which inter-experimental variability has been removed to a varying degree. Usually we have $\dim(\mathcal{Z}) \ll \dim(\mathcal{X})$ and so, for different downstream evaluations, we map between these representations: (*i*) with least squares reconstructions ($\mathcal{Z} \to \mathcal{X}$), or (*ii*) principle component projections ($\mathcal{X} \to \mathcal{Z}$) (see Appendix).

4.2.1 Unsupervised Model

We parameterize the functions f, g and h (Figure 1) with neural networks. In the unsupervised model, the function $g : \mathcal{Z} \to \mathcal{X}$ provides a reconstruction $\hat{x} := g(z),\ \hat{x} \in \mathcal{X}$ of the data. With the concurrent task (eq. 3) of minimizing the predictability of the domain D, this reconstruction should only contain parts of the original data that are indiscernible across experiments. Since the purpose of our method is to `Remove`, `Adversarially`, `Variability` from datasets collected in different `Experiments`, we term our model `RAVE`.

4.2.2 Supervised Model

In the supervised setting, we have partial knowledge about the biological signal S. The function $g : \mathcal{Z} \to \mathcal{X} \times \mathcal{S}$ now returns a reconstruction as well as a prediction of that signal $(\hat{x}, \hat{s}) := g(z),\ \hat{s} \in \mathcal{S}$. When optimizing equation (4), this additional task is equivalent to discriminative adversarial domain adaptation [37]. In the particular data that we work with, we have two datasets $\mathcal{D} := \{A, B\}$, but the biological signal S consists of cell type labels which are only available in the first dataset A. Thus, more accurately, this presents a *semi-supervised* scenario where one wishes to classify a newly recorded dataset according to some existing classification scheme. We term this extended version of our model `RAVE+`.

4.2.3 Training and Optimization Details

All of our models are implemented and optimized in PyTorch [43]. For both `RAVE` and `RAVE+`, we use the same model architecture, they only differ in the objective function. We randomly split the data

into training, validation and test set and train all models with empirical risk minimization. Model weights are trained with stochastic gradient descent using one instance of the Adam optimizer [44] for the outer minimization of f and g in equations (3) and (4); and then a second instance of Adam for the inner minimization of h in those equations. We optimize hyperparameters through random search [45] on the validation set and report performances on the test set which is only used for final evaluation. In the random search, we test different learning rates for both optimizers, and also different training schedules. We additionally search over depth, width and drop-out rate for each of the neural networks (f, g, h), as well as the trade-off parameter λ introduced in equation (3). Finally, we explore training the inner optimization (h, estimating $I(Z; D)$) more often than the outer optimization, which proved more stable and effective in early experiments.

4.2.4 Comparison Models

We test three different methods for comparison with our model. Our simplest comparison model (*Linear*) is a linear model that projects out the contrast between the dataset indicator variables (see Appendix). This has an analytic solution and no hyperparameters, serving as a baseline to get an estimate of the correction quality achieved by a standard method in classical statistics. The other two methods (*scGen* and *Harmony*) are run in an unsupervised learning mode without cell type information. Even though *scGen* could be utilized to run in a supervised mode with cell type information, this is not specified in a semi-supervised setting with only partial cell type labels available.

4.3 Performance Evaluation

For evaluating the correction performed by the various methods, we analyze their output with respect to dataset-mixing (achieved by removing inter-experimental variability) and preservation of signal information.

4.3.1 Dataset-Mixing

The Rand index [46] measures similarity between two clusterings; the adjusted Rand index (ARI) is the Rand index adjusted for chance level (see Appendix) which was recently used by Tran et al. [28] to assess the quality of dataset-mixing in genomics. It takes as input the true and the predicted labels for a set of samples. We define $\text{ARI}_{dom}(z) := \text{ARI}(d, \hat{d}_z)$ with d the original domain labels (of the test set) and $\hat{d}_z$ the domain labels predicted by a classifier trained on z. On the raw data ($z = x$), we expect ARI_{dom} to be high due to inter-experimental variability. After successful correction (with z the output of a model), we expect ARI_{dom} to be low indicating good dataset-mixing.

In addition, we compute the accuracy (Acc_{dom}) of a domain classifier with the objective to predict the domain labels based on the input data. For low dataset-mixing, we expect a high Acc_{dom} as it should be trivial for the classifier to differentiate the datasets. However, after removing inter-experimental variability, Acc_{dom} is supposed to be close to chance level ($\sim 64\%$, cf. Table 1), which would indicate successful dataset-mixing. For the domain classifier, we use a random forest classifier with cross-validated hyperparameters for each model (see Appendix). This is crucial, because a powerful encoder f might hide (through multiple nonlinear transformations) domain information from a simple classifier, but still recover that information in an equally powerful decoder g. Conversely, we observe that overly expressive random forest classifiers, tend to overfit on the training set, thus underestimating the preserved domain or type information on the test set.

4.3.2 Preservation of Signal Information

In the unsupervised setting, to assess the amount of information preserved about the original data x during the process of removing inter-experimental variability, we evaluate the rank correlation $\text{Corr}(x, \hat{x})$ between input x and reconstruction $\hat{x}$. In the (semi-) supervised setting, we have reference cell type labels s_A for dataset A. To estimate how much of this information is preserved, we predict cell type labels $\hat{s}_A$ from z_A with random forest classifiers like above (see Figure 2; Appendix for further details).

If a method succeeds at preserving signal information in z after removing inter-experimental variability, then we expect the classifier to have a high accuracy (Acc_{type}). Deteriorating classification

performance between predicting $\hat{s}_A$ from raw data x_A versus predicting it from the model output z_A would indicate signal loss.

Additionally, we would like to evaluate how well cell types can be distinguished and how biologically plausible they are for the unlabeled dataset *B*. To this end, we apply the classifiers to predict cell type labels $\hat{s}_B$ from z_B. One direct comparison is between the distributions over cell types in $\hat{s}_B$ and as expected from electron-microscopy (EM) data [47–50] (Figure 5A). However, we can also evaluate the accuracy of these predictions by making use of BC axonal stratification profiles obtained from the same EM data. From those data, we know where in the IPL a BC type stratifies its axon terminals. Thus, we can compare the distribution over IPL depth for the predicted cell types ($\hat{s}_B$) with the distributions expected from EM data. We quantify the difference between the expected and predicted distributions by calculating the Jensen-Shannon distance. We define the depth score (DS) as the mean Jensen-Shannon distance between those two distributions (Table 2). Additionally, we evaluate the robustness of cell type labels $\hat{s}_B$ by fitting the classifier ten times with different seeds and calculating the average ARI between different runs, giving the ARI_{type} score (Table 1).

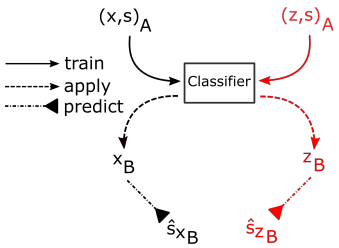

Figure 2: **Workflow.** Evaluating the preservation of signal information: A classifier gets trained on the labeled dataset *A* (either $(x, s)_A$ or $(z, s)_A$) and applied to dataset *B* to predict labels (either $\hat{s}_{x_B}$ or $\hat{s}_{z_B}$). The predicted labels are then used for further evaluation.

5 Results

5.1 Simulation Experiments

First of all, we validated that our model performs as expected on simulated ground truth data. To do this, we generated bipolar cell responses for all 14 cell types based on the published bipolar cell model in Schröder et al. [51]. To simulate different individual neurons, we added small perturbations to the model weights for each cell type until we matched the intra-cell-type variability observed in the real data. Thus, we generated $N = 1000$ distinct neurons for each of the 14 BC types. Approximating the differences of the two datasets in the paper, we presented the model with the slightly altered versions of the stimulus from the actual experiments (see Appendix B). We additionally added white noise to match realistic signal to noise ratios, as estimated from repeated stimulations of the real neurons. This resulted in two datasets ('A' and 'B') with similar intra- and inter-experimental variability as observed in the real data, but with known ground truth cell type labels.

The results are discussed here and, additionally, they are presented in Appendix Fig. 9. We first confirmed that the classifiers are indeed perfectly able to separate these two artificial datasets based on their systematic differences (domain accuracy on raw simulated data: 1.0). However, cell type classifiers trained on dataset A fail completely on dataset B indicating severe inter-experimental variability and a failure to transfer models across datasets (type accuracy on dataset A: 0.98, type accuracy on dataset B: 0.16). In contrast, after correction with RAVE+, the performance of a classifier trained to distinguish the two datasets drops from 1.0 to 0.66 (chance level 0.5), indicating strong removal of inter-experimental variability. Importantly, we find that a classifier trained on the output of RAVE+ on dataset A does now generalize to dataset B and recovers the ground truth cell labels nearly perfectly (type accuracy 0.99). This constitutes an important validation of our model.

5.2 Unsupervised Removal of Inter-Experimental Variability

All methods tested (*Linear*, *Harmony*, *scGen*, `RAVE` and `RAVE+`) succeed at retaining a significant amount of information about x in $\hat{x}$, reflected by high correlations between x and $\hat{x}$ (Table 1). $\text{Corr}(x, \hat{x})$ reaches similar levels for data from both datasets, suggesting that both datasets are modified to find a midway representation. This impression is confirmed when visualizing x_A, x_B and $\hat{x}_A$ and $\hat{x}_B$ next to each other (Figure 3). Moreover, we note that a side-effect of the alignment is a more general denoising that `RAVE` tends to perform along with removing inter-experimental variability. We recognize that this a desirable feature that we will study further in future research.

Model	$\text{Corr}_A \uparrow$	$\text{Corr}_B \uparrow$	$\text{Acc}_{dom} \downarrow$	$\text{Acc}_{type} \uparrow$	$\text{ARI}_{dom} \downarrow$	$\text{ARI}_{type} \uparrow$
Raw	100	100	99.8 (0.1)	77.4 (0.9)	99.3 (0.4)	37.3 (3.3)
Linear	**99.0** (0.5)	**97.0** (1.7)	99.5 (0.2)	83.4 (0.8)	98.1 (0.7)	7.6 (1.7)
Harmony	72.0 (10.7)	72.0 (13.8)	94.2 (0.4)	82.5 (0.5)	78.0 (1.6)	31.4 (2.3)
scGen	78.0 (9.8)	80.0 (10.5)	99.6 (0.1)	**84.7** (0.8)	98.7 (0.3)	14.3 (2.6)
RAVE	60.0 (12.8)	58.0 (17.3)	77.5 (0.5)	69.5 (0.4)	28.9 (1.2)	81.2 (2.7)
RAVE+	59.0 (14.8)	58.0 (19.1)	**65.9** (0.9)	78.6 (0.8)	**10.0** (1.2)	**83.7** (2.3)

Table 1: **Model Comparison.** All entries in percentage. Mean and standard deviation metric scores across 10 random seeds. Bold font in each row indicates best score. Corr_A (Corr_B) is the correlation of corrected data from dataset *A* (B) with its raw data. Acc_{dom} (Acc_{type}) is the accuracy of the domain (cell type) classifier. For ARI_{dom} and ARI_{type} see Section 4.3.

We show mean traces for exemplary cell types from dataset *A*, and mean traces of cells from dataset *B* whose cell type labels we predict twice, first based on x (left pathway in Figure 2) and then again based on $\hat{x}$ (right pathway in Figure 2, but on $\hat{x}_{\texttt{RAVE}}$ instead of z). As expected, inter-experimental variability obscuring the common signal s behind x_A and x_B causes the cell type assignment to fail; the similarity between responses of cells assigned to the same cell type, but coming from the different datasets is low (Figure 3, BC type 5t). Repeating the classification pipeline based on x with the same classifier architecture and different seeds yields highly variable cell type predictions for dataset *B* (Table 1, ARI_{type}) despite high prediction accuracy on dataset *A* (Table 1, Acc_{type}). This demonstrates a failure in transferring to dataset *B*, and not the classification itself. These results on the raw data x affected by inter-experimental variability were expected; however, the same pattern - low ARI_{type} (dataset *B*) and high Acc_{type} (dataset *A*) - is observed for *Harmony*, *scGen* and the *Linear* model. This suggests that these methods fail at removing inter-experimental variability. The high domain accuracy achieved by a classifier trained on the outputs of these models confirms this conclusion. RAVE, on the other hand, succeeds at significantly lowering domain accuracy Acc_{dom}, while at the same time maintaining high scores for ARI_{type} and Acc_{type}.

5.3 Supervised Removal of Inter-Experimental Variability

RAVE+ extends RAVE to the (semi-) supervised setting where (partial) signal information is present. RAVE+ excels at removing inter-experimental variability (Table 1, Acc_{dom} and ARI_{dom}) and at the same time retaining signal information (Table 1, Acc_{type} and ARI_{type}). A low dimensional t-SNE

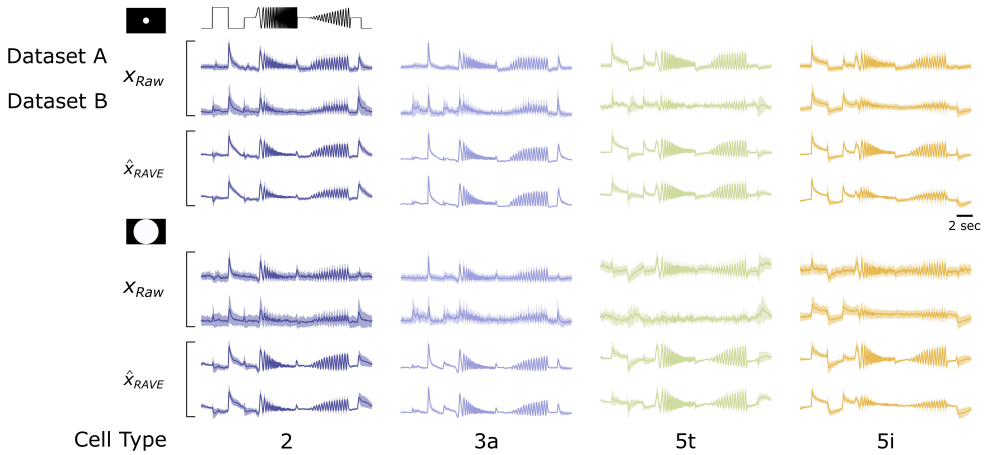

Figure 3: **Exemplary Cell Type Responses from both Datasets to the Chirp Stimuli.** Four bipolar cell type responses of the types 2, 3a, 5t and 5i to the local (top panel) and full-field (bottom panel) chirp of raw data x_{Raw} (two top upper panels) and reconstructed data $\hat{x}_{\texttt{RAVE}}$ (two bottom lower panels) by RAVE for both datasets *A* and *B*. Each column shows the mean responses of one cell type (standard deviation shaded).

[52] embedding (Figure 4) shows that cells from datasets *A* and *B* are mapped onto the same cell type "islands". The distribution of types across IPL depth predicted by a classifier trained on $z_{\texttt{RAVE+}}$ matches the expected anatomical distributions better than for all other methods (Figure 5 and Table 2). This provides a valuable validation of the estimate $\hat{s}_B$ learned by RAVE+ in the absence of ground truth knowledge of s_B.

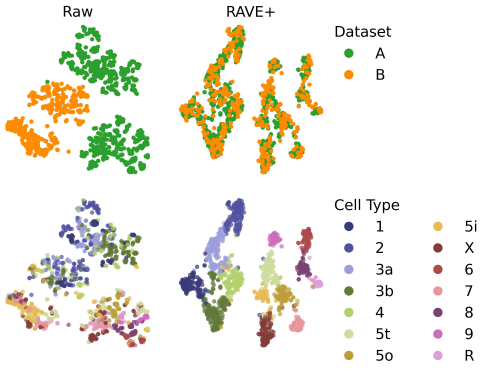

Figure 4: **Dataset Embeddings**. t-SNE embeddings of the test set of raw (left column) and corrected output data by RAVE+ (right column). Embedded cells are color-coded by dataset (top row) and cell type (bottom row). Cell type labels for the raw data of dataset *B* (bottom left) were predicted using a cell type classifier trained on the raw data of dataset *A* (Figure 2, left pathway).

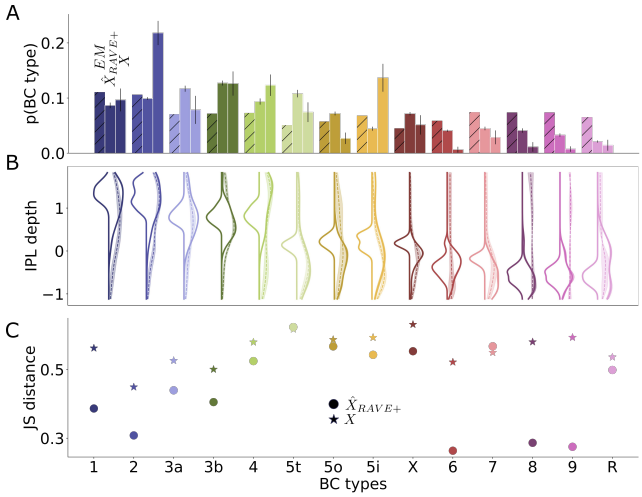

Figure 5: **Distribution Across BC Types and IPL Depths.** We compare the expected and predicted distribution of BCs from dataset *B* across the 14 types and across IPL depth. **(A)** Probability that a BC belongs to a certain type as estimated from EM data; as estimated from BC type labels predicted on $\hat{x}_{\texttt{RAVE+}}$; and as estimated from BC type labels predicted on x. Error bars indicate SD across 10 seeds of the classifier. (**B**) Distributions per cell type over IPL depth for EM data (distribution shown to the left), RAVE+ output (solid line to the right) and raw data (dashed line to the right). Shaded area around the distributions shown to the right indicate SD across 10 seeds of the classifier. (**C**) JS distances corresponding to the distributions in B).

5.4 Downstream Analyses on Reconstructed Traces

As in our unsupervised setting, it is common that no particular signal information is available and that one wants to remove inter-experimental variability from the data to perform further downstream

BC Type	1	2	3a	3b	4	5t	5o	5i	X	6	7	8	9	R	∅	all
Raw	**34**	52	**38**	42	56	63	62	54	60	52	**51**	38	58	**37**	50	31
Linear	58	40	53	51	57	62	57	56	62	55	57	58	58	53	56	34
Harmony	42	32	42	47	54	59	58	56	61	37	59	**26**	58	38	48	23
scGen	51	38	48	48	56	59	**56**	59	63	55	57	56	56	42	53	31
RAVE	41	32	**38**	**40**	57	**58**	62	55	66	55	55	38	50	50	50	23
RAVE+	38	**30**	43	**40**	**52**	62	**56**	**54**	**55**	**26**	56	28	**27**	49	**44**	**17**

Table 2: **Depth Score Comparison.** All entries in percentage, lower is better. Bold font in each row indicates best score. Depth Score - Jensen-Shannon (JS) distance between predicted types and EM depth distribution: $JS(p_{EM}(depth|type = t), p_{model}(depth|type = t))$. Last column ("all"): $JS(p_{EM}(type), p_{model}(type))$.

analyses. We show that a previously demonstrated biological effect, obscured by inter-experimental variability in x, emerges when performing the same analyses on the reconstructed traces $\hat{x}$ obtained from RAVE. Full-field visual stimulation has been shown to decorrelate responses from different BC types compared to local stimulation due to inhibitory feedback from amacrine cells (see Figure 3A, B in [2]). We expect this fundamental feature to be present in dataset *B*, but cannot fully reproduce it if we assign cells of dataset *B* to cell types based on the raw data (Figure 2, left pathway; and Figure 6A). However, using the reconstructed traces $\hat{x}_{\texttt{RAVE}}$, the expected feature is unmasked (Figure 2, right pathway, but on $\hat{x}_{\texttt{RAVE}}$ instead of z; and Figure 6B). Here, the mean responses to the local chirp are more correlated across cell types than the full-field responses (Figure 6B, left panel). This can also be seen when comparing the mean correlations between local and full-field chirp responses for each cell type with all other cell types, both of the same and the opposite response polarity (On and Off polarity) (Figure 6B, right panel).

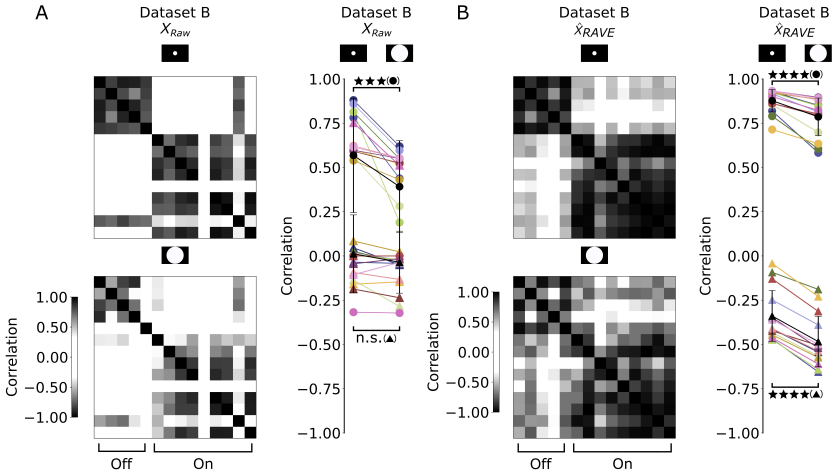

Figure 6: **Removing Inter-Experimental Variability Reveals Biological Feature.** (**A**) Correlation matrices show the correlations between mean responses per cell type to local (top) and full-field (bottom) chirp of raw data x from dataset *B*. The right panel represents the mean correlation for each cell type mean response with all other types of the same (circle) and opposite (triangle) response polarity between local and full-field chirp shown for raw responses. (x: mean correlation same polarity: p_{local} = 0.57 and $p_{full\text{-}field}$ = 0.39, P < 0.005; opposite polarity: p_{local} = 0.01 and $p_{full\text{-}field}$ = -0.04, n.s.; n = 14, non-parametric paired Wilcoxon signed-rank test).(**B**) Same analysis as **A**, but with the reconstructed responses obtained from RAVE. ($\hat{x}_{\texttt{RAVE}}$: mean correlation same polarity: p_{local} = 0.88 and $p_{full\text{-}field}$ = 0.79, P < 0.0005; opposite polarity: p_{local} = -0.34 and $p_{full\text{-}field}$ = -0.48, P < 0.0005; n = 14, non-parametric paired Wilcoxon signed-rank test).

6 Limitations

Our method is limited to datasets where neurons were presented with the same stimulus. For other kinds of data, such as neural recordings from free behavioral paradigms where each trial will be different, it will be difficult to 'align' neural responses in a meaningful way. One solution to this could be to learn a shared embedding space [see 53], from which domain effects are removed, but distinct encoders f_i and decoders g_i for different trials i. In another setting, where different stimuli are presented between experiments, one might resort to an approach like Shah et al. [3]. Nevertheless, we do acknowledge that the data in our applications consists of *ex vivo* retinal recordings which have little to no attentional effects or task-dependent noise correlations like they would be present in *in vivo* cortical data. We are optimistic that our framework of adversarially removing inter-experimental variability is still a promising approach in those settings, under the constraint that a much more severe trade-off may need to be made between retaining signal and removing domain shifts.

7 Discussion

We present a framework to remove inter-experimental variability from functional recordings in systems neuroscience. To the best of our knowledge, this is the first application of domain adaptation methods to this kind of data. Using our unsupervised (`RAVE`) and (semi-) supervised (`RAVE+`) approaches, we demonstrate that we are able to remove inter-experimental variability while retaining signal information, which allows us to robustly predict cell type labels for a new dataset. We validate those predictions using an anatomy-based comparison to existing EM data.

Furthermore, our unsupervised approach `RAVE` is able to remove inter-experimental variability without cell type information. By using the corrected dataset B, we unmask biological effects, obscured by inter-experimental variability, that have been previously described for dataset A. Thus, by allowing the integration and alignment of functional recordings across experiments, we show that biological effects in the data become more pronounced when using our model approaches. Inter-experimental variability is ubiquitous and we hope that this method will become a helpful resource to many experimenters as we make the code toolbox publicly available.

We believe that our method can also make a contribution to systems neuroscience research in the context of the 3Rs (Replacement, Reduction and Refinement) for animal ethics: By enabling detection of more subtle biological signals after removal of inter-experimental variability, fewer animals may be needed to test a specific hypothesis. Lastly, we acknowledge that the removal of inter-experimental variability from any kind of data (thus not only within systems neuroscience) can be useful in various applications. Virtually any analysis that aggregates data across experiments can be confounded by inter-experimental variability. Consequently, we cannot exclude the possibility that some military application will find value in this approach. Although unlikely, we cannot fully anticipate such developments. Therefore we condemn, without any exceptions, the use of `RAVE(+)` for any warlike applications or other nefarious purposes.

8 Acknowledgments and Disclosure of Funding

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) project number 335549539/GRK2381 and the CRC 1233 "Robust Vision" (grant number 276693517). Moreover, this work was partially supported by a Research Council of Norway FRIPRO grant (90532703). PB is a member of the Machine Learning Cluster of Excellence, EXC number 2064/1 – Project number 390727645 and the Tübingen AI Center (FKZ: 01IS18039A). He was supported through a Heisenberg Professorship by the DFG (BE5601/8-1).

References

[1] Tom Baden, Philipp Berens, Katrin Franke, Miroslav Román Rosón, Matthias Bethge, and Thomas Euler. The functional diversity of retinal ganglion cells in the mouse. *Nature*, 529(7586):345–350, 2016.

[2] Katrin Franke, Philipp Berens, Timm Schubert, Matthias Bethge, Thomas Euler, and Tom Baden. Inhibition decorrelates visual feature representations in the inner retina. *Nature*, 542(7642):439–444, 2017.

[3] Nishal Shah, Nora Brackbill, Ryan Samarakoon, Colleen Rhoades, Alexandra Kling, Alexander Sher, Alan Litke, Yoram Singer, Jonathon Shlens, and EJ Chichilnisky. Individual variability of neural computations in the primate retina. *bioRxiv*, pages 1–22, 2021.

[4] Blake A Richards, Timothy P Lillicrap, Philippe Beaudoin, Yoshua Bengio, Rafal Bogacz, Amelia Christensen, Claudia Clopath, Rui Ponte Costa, Archy de Berker, Surya Ganguli, et al. A deep learning framework for neuroscience. *Nature neuroscience*, 22(11):1761–1770, 2019.

[5] Zhijian Zhao, David A. Klindt, André Maia Chagas, Klaudia P. Szatko, Luke Rogerson, Dario A. Protti, Christian Behrens, Deniz Dalkara, Timm Schubert, Matthias Bethge, Katrin Franke, Philipp Berens, Alexander S. Ecker, and Thomas Euler. The temporal structure of the inner retina at a single glance. *Scientific Reports*, 10(1):1–17, 2020.

[6] Stephanie C Hicks, F William Townes, Mingxiang Teng, and Rafael A Irizarry. Missing data and technical variability in single-cell rna-sequencing experiments. *Biostatistics*, 19(4):562–578, 2018.

[7] Jing Jiang. A literature survey on domain adaptation of statistical classifiers. *URL: http://sifaka. cs. uiuc. edu/jiang4/domainadaptation/survey*, 3:1–12, 2008.

[8] John Blitzer, Koby Crammer, Alex Kulesza, Fernando Pereira, and Jennifer Wortman. Learning bounds for domain adaptation. 2008.

[9] Baochen Sun, Jiashi Feng, and Kate Saenko. Return of frustratingly easy domain adaptation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30, 2016.

[10] Dario Amodei, Chris Olah, Jacob Steinhardt, Paul Christiano, John Schulman, and Dan Mané. Concrete problems in ai safety. *arXiv preprint arXiv:1606.06565*, 2016.

[11] Shai Ben-David, John Blitzer, Koby Crammer, Alex Kulesza, Fernando Pereira, and Jennifer Wortman Vaughan. A theory of learning from different domains. *Machine learning*, 79(1):151–175, 2010.

[12] 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, et al. Discovering precise temporal patterns in large-scale neural recordings through robust and interpretable time warping. *Neuron*, 105(2):246–259, 2020.

[13] Fabian Sinz, Alexander S Ecker, Paul Fahey, Edgar Walker, Erick Cobos, Emmanouil Froudarakis, Dimitri Yatsenko, Zachary Pitkow, Jacob Reimer, and Andreas Tolias. Stimulus domain transfer in recurrent models for large scale cortical population prediction on video. *Advances in Neural Information Processing Systems*, 31:7199–7210, 2018.

[14] Oleksandr Sorochynskyi, Stéphane Deny, Olivier Marre, and Ulisse Ferrari. Predicting synchronous firing of large neural populations from sequential recordings. *PLoS computational biology*, 17(1):e1008501, 2021.

[15] Alexander S Ecker, Philipp Berens, R James Cotton, Manivannan Subramaniyan, George H Denfield, Cathryn R Cadwell, Stelios M Smirnakis, Matthias Bethge, and Andreas S Tolias. State dependence of noise correlations in macaque primary visual cortex. *Neuron*, 82(1):235–248, 2014.

[16] Jonathan Jouty, Gerrit Hilgen, Evelyne Sernagor, and Matthias H Hennig. Non-parametric physiological classification of retinal ganglion cells in the mouse retina. *Frontiers in Cellular Neuroscience*, 12:481, 2018.

[17] Gabriele Beate Schweikert, Christian Widmer, Bernhard Schölkopf, and Gunnar Rätsch. An empirical analysis of domain adaptation algorithms for genomic sequence analysis. In *NIPS*, volume 8, pages 1433–1440, 2008.

[18] Jonathan Bryan Dayton. Adversarial deep neural networks effectively remove nonlinear batch effects from gene-expression data. 2019.

[19] Tongxin Wang, Travis S Johnson, Wei Shao, Zixiao Lu, Bryan R Helm, Jie Zhang, and Kun Huang. Bermuda: a novel deep transfer learning method for single-cell rna sequencing batch correction reveals hidden high-resolution cellular subtypes. *Genome biology*, 20(1):1–15, 2019.

[20] Jingshu Wang, Divyansh Agarwal, Mo Huang, Gang Hu, Zilu Zhou, Chengzhong Ye, and Nancy R Zhang. Data denoising with transfer learning in single-cell transcriptomics. *Nature methods*, 16(9):875–878, 2019.

[21] Dongfang Wang, Siyu Hou, Lei Zhang, Xiliang Wang, Baolin Liu, and Zemin Zhang. imap: integration of multiple single-cell datasets by adversarial paired transfer networks. *Genome biology*, 22(1):1–24, 2021.

[22] Xiangjie Li, Kui Wang, Yafei Lyu, Huize Pan, Jingxiao Zhang, Dwight Stambolian, Katalin Susztak, Muredach P Reilly, Gang Hu, and Mingyao Li. Deep learning enables accurate clustering with batch effect removal in single-cell rna-seq analysis. *Nature communications*, 11(1):1–14, 2020.

[23] Eugene Lin, Sudipto Mukherjee, and Sreeram Kannan. A deep adversarial variational autoencoder model for dimensionality reduction in single-cell rna sequencing analysis. *BMC bioinformatics*, 21(1):1–11, 2020.

[24] Gökcen Eraslan, Lukas M Simon, Maria Mircea, Nikola S Mueller, and Fabian J Theis. Single-cell rna-seq denoising using a deep count autoencoder. *Nature communications*, 10(1):1–14, 2019.

[25] Ilya Korsunsky, Nghia Millard, Jean Fan, Kamil Slowikowski, Fan Zhang, Kevin Wei, Yuriy Baglaenko, Michael Brenner, Po-ru Loh, and Soumya Raychaudhuri. Fast, sensitive and accurate integration of single-cell data with harmony. *Nature methods*, 16(12):1289–1296, 2019.

[26] Mohammad Lotfollahi, F Alexander Wolf, and Fabian J Theis. scgen predicts single-cell perturbation responses. *Nature methods*, 16(8):715–721, 2019.

[27] Malte D Luecken, Maren Buttner, Kridsadakorn Chaichoompu, Anna Danese, Marta Interlandi, Michaela F Müller, Daniel C Strobl, Luke Zappia, Martin Dugas, Maria Colomé-Tatché, et al. Benchmarking atlas-level data integration in single-cell genomics. *BioRxiv*, 2020.

[28] Hoa Thi Nhu Tran, Kok Siong Ang, Marion Chevrier, Xiaomeng Zhang, Nicole Yee Shin Lee, Michelle Goh, and Jinmiao Chen. A benchmark of batch-effect correction methods for single-cell rna sequencing data. *Genome biology*, 21(1):1–32, 2020.

[29] Rudy Moddemeijer. On estimation of entropy and mutual information of continuous distributions. *Signal processing*, 16(3):233–248, 1989.

[30] Liam Paninski. Estimation of entropy and mutual information. *Neural computation*, 15(6):1191–1253, 2003.

[31] Janett Walters-Williams and Yan Li. Estimation of mutual information: A survey. In *International Conference on Rough Sets and Knowledge Technology*, pages 389–396. Springer, 2009.

[32] Ben Poole, Sherjil Ozair, Aaron Van Den Oord, Alex Alemi, and George Tucker. On variational bounds of mutual information. In *International Conference on Machine Learning*, pages 5171–5180. PMLR, 2019.

[33] Izrail Moiseevich Gel'fand and Akiva Moiseevich Yaglom. Computation of the amount of information about a stochastic function contained in another such function. *Uspekhi Matematicheskikh Nauk*, 12(1):3–52, 1957.

[34] Jörn-Henrik Jacobsen, Jens Behrmann, Richard Zemel, and Matthias Bethge. Excessive invariance causes adversarial vulnerability. *arXiv preprint arXiv:1811.00401*, 2018.

[35] David Barber and Felix V Agakov. Information maximization in noisy channels: A variational approach. *Advances in Neural Information Processing Systems*, 16:201–208, 2003.

[36] Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. *arXiv preprint arXiv:1406.2661*, 2014.

[37] Eric Tzeng, Judy Hoffman, Kate Saenko, and Trevor Darrell. Adversarial discriminative domain adaptation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7167–7176, 2017.

[38] Winfried Denk, James H Strickler, and Watt W Webb. Two-photon laser scanning fluorescence microscopy. *Science*, 248(4951):73–76, 1990.

[39] Winfried Denk and Peter B Detwiler. Optical recording of light-evoked calcium signals in the functionally intact retina. *Proceedings of the National Academy of Sciences*, 96(12):7035–7040, 1999.

[40] Thomas Euler, Susanne E Hausselt, David J Margolis, Tobias Breuninger, Xavier Castell, Peter B Detwiler, and Winfried Denk. Eyecup scope—optical recordings of light stimulus-evoked fluorescence signals in the retina. *Pflügers Archiv-European Journal of Physiology*, 457(6):1393–1414, 2009.

[41] Thomas Euler, Silke Haverkamp, Timm Schubert, and Tom Baden. Retinal bipolar cells: elementary building blocks of vision. *Nature Reviews Neuroscience*, 15(8):507–519, 2014.

[42] Jonathan S. Marvin, Bart G. Borghuis, Lin Tian, Joseph Cichon, Mark T. Harnett, Jasper Akerboom, Andrew Gordus, Sabine L. Renninger, Tsai Wen Chen, Cornelia I. Bargmann, Michael B. Orger, Eric R. Schreiter, Jonathan B. Demb, Wen Biao Gan, S. Andrew Hires, and Loren L. Looger. An optimized fluorescent probe for visualizing glutamate neurotransmission. *Nature Methods*, 10(2):162–170, 2013.

[43] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *arXiv preprint arXiv:1912.01703*, 2019.

[44] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[45] James Bergstra and Yoshua Bengio. Random search for hyper-parameter optimization. *Journal of machine learning research*, 13(2), 2012.

[46] William M Rand. Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical association*, 66(336):846–850, 1971.

[47] Moritz Helmstaedter, Kevin L Briggman, Srinivas C Turaga, Viren Jain, H Sebastian Seung, and Winfried Denk. Connectomic reconstruction of the inner plexiform layer in the mouse retina. *Nature*, 500(7461):168–174, 2013.

[48] Christian Behrens, Timm Schubert, Silke Haverkamp, Thomas Euler, and Philipp Berens. Connectivity map of bipolar cells and photoreceptors in the mouse retina. *Elife*, 5:e20041, 2016.

[49] Jinseop S. Kim, Matthew J. Greene, Aleksandar Zlateski, Kisuk Lee, Mark Richardson, Srinivas C. Turaga, Michael Purcaro, Matthew Balkam, Amy Robinson, Bardia F. Behabadi, Michael Campos, Winfried Denk, and H. Sebastian Seung. Space-time wiring specificity supports direction selectivity in the retina. *Nature*, 509(7500):331–336, 2014.

[50] Matthew J. Greene, Jinseop S. Kim, and H. Sebastian Seung. Analogous Convergence of Sustained and Transient Inputs in Parallel On and Off Pathways for Retinal Motion Computation. *Cell Reports*, 14(8):1892–1900, 2016.

[51] Cornelius Schröder, David A. Klindt, Sarah Strauss, Katrin Franke, Matthias Bethge, Thomas Euler, and Philipp Berens. System identification with biophysical constraints: A circuit model of the inner retina. *bioRxiv*, 2020.

[52] Geoffrey Hinton and Sam T Roweis. Stochastic neighbor embedding. In *NIPS*, volume 15, pages 833–840. Citeseer, 2002.

[53] Rohan Gala, Nathan Gouwens, Zizhen Yao, Agata Budzillo, Osnat Penn, Bosiljka Tasic, Gabe Murphy, Hongkui Zeng, and Uygar Sümbül. A coupled autoencoder approach for multi-modal analysis of cell types. *arXiv preprint arXiv:1911.05663*, 2019.

[54] Klaudia P Szatko, Maria M Korympidou, Yanli Ran, Philipp Berens, Deniz Dalkara, Timm Schubert, Thomas Euler, and Katrin Franke. Neural circuits in the mouse retina support color vision in the upper visual field. *Nature communications*, 11(1):1–14, 2020.

[55] DI Vaney. 'coronate' amacrine cells in the rabbit retina have the 'starburst'dendritic morphology. *Proceedings of the Royal society of London. Series B. Biological sciences*, 220(1221):501–508, 1984.

[56] Katrin Franke, André Maia Chagas, Zhijian Zhao, Maxime J.Y. Zimmermann, Philipp Bartel, Yongrong Qiu, Klaudia P. Szatko, Tom Baden, and Thomas Euler. An arbitrary-spectrum spatial visual stimulator for vision research. *eLife*, 8:1–28, 2019.

[57] F Alexander Wolf, Philipp Angerer, and Fabian J Theis. Scanpy: large-scale single-cell gene expression data analysis. *Genome biology*, 19(1):1–5, 2018.

[58] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.

[59] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830, 2011.

[60] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[61] Pavlin G Poličar, Martin Stražar, and Blaž Zupan. opentsne: a modular python library for t-sne dimensionality reduction and embedding. *BioRxiv*, page 731877, 2019.

Appendix

A Model details

A.1 Linear mappings between z and x

Usually, we have data $x \in \mathbb{R}^{N \times D_1}$ and latent representation $z \in \mathbb{R}^{N \times D_2}$ with N the number of neurons, D_1 the dimensionality of the data, D_2 the dimensionality of the latent space and, usually, $D_1 \gg D_2$. In cases where a method m does only produce some latent representation z_m, we fit a reconstruction $\hat{x}_m = W z_m$ with a least squares projection $W = (z_m^T z_m)^{-1} z_m^T x$. In cases where a method m does only produce some reconstruction $\hat{x}_m$, we produce a simple latent representation z_m by extracting the first D_2 columns of the left singular vectors U from the singular value decomposition $x = USV^T$. Both of these projections are fitted on the training data, then fixed and also used on the validation and test data.

B Data

We used three datasets, where the first two (dataset *A* [2] n=8417 cells; *B* [54] n=4600) are two-photon recordings of mouse retinal bipolar cell (BC) responses to the chirp stimuli (local and full-field, see [2] for details). Both datasets were used for model fitting and removal of inter-experimental variability. For the validation of cell type predictions made by the different models, we used the third dataset, which comprises EM data of axonal stratification profiles as probability distribution of each BC type [47–50].

The inter-experimental variability between the two functional datasets may originate from, at least, the three following differences between the datasets: (*i*) dataset *A* recorded BCs mostly at certain IPL depths ('ChAT-bands', which are landmarks within the IPL [55]) using tangential scans parallel to the retinal layers, whereas dataset *B* used axial scans employing an electrically tunable lens to record from BCs across the entire IPL simultaneously [5], resulting in different sampling distributions; (*ii*) the chirp stimulus used in dataset *B* differs slightly as the sinusoidal intensity modulation of the increasing frequency is marginally slower; (*iii*) dataset *A* did not employ a gamma correction of the display device to linearize its intensity curve, resulting in slightly different stimulus contrasts [56].

C Training Results

The outcome of the random search can be seen in Figure 7, showing metrics on the validation set for both models. To select the best `RAVE` model, we picked the point in the top right corner (center plot, first row, Figure 7). This was the model with the highest $I(Z; X)$, i.e. correlation, and the lowest $I(Z; D)$, i.e. domain classification accuracy. To select the best `RAVE+` model, we picked the (`RAVE+`) point in the top right corner of the $3D$ space spanned by $\{I(Z; X), I(Z; S), -I(Z; D)\}$, i.e. the model with the best reconstruction and cell type prediction accuracy but with the lowest domain prediction accuracy.

Moreover, Figure 7 also demonstrates the trade-off between maximizing $I(Z; X)$ and $I(Z; S)$ and minimizing $I(Z; D)$. In the top row on the left, one can see that models with high $I(Z; X)$ also tend to have a high $I(Z; S)$, indicating that these two tasks can be performed well at the same time (this is what we mean by 'synergy' in the title; naturally, we cannot make a causal statement here). In the top row middle, one can see for models that achieve a high $I(Z; X)$ (some hyperparameter configurations in the random search simple lead to bad models), that there is a negative slope with respect to $I(Z; D)$, indicating that there is a trade-off between optimizing these two objectives. The same can be seen in the top row on the right with respect to $I(Z; S)$ and $I(Z; D)$. The bottom row of Figure 7 zooms in on the high performing models (see axes limits) and indicates the rank correlations. As stated above, we find a positive correlation between $I(Z; X)$ and $I(Z; S)$ (i.e. no conflict), but a negative correlation between $I(Z; X)$ and $I(Z; D)$, and between $I(Z; S)$ and $I(Z; D)$ (i.e. a trade-off).

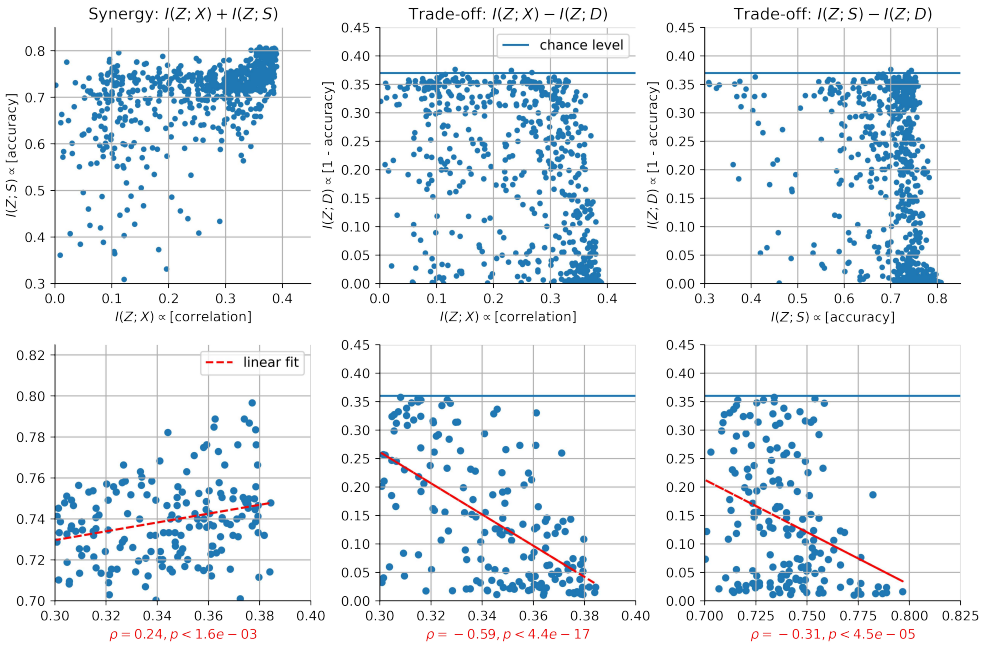

Figure 7: **Random Search Result**. Optimizing models with different hyperparameters shows how the terms in the objective function interact. The top row shows all models, the bottom row only filtered (high-performing) models within the indicated axes ranges. The red lines in the bottom plot indicate linear fits and the red axis labels show the rank correlation coefficients ρ and p values.

D Details for Comparison Models

D.1 Linear Model

Let our full dataset $x \in \mathbb{R}^{(N+M)\times D}$ consists of the concatenated datasets $x_A \in \mathbb{R}^{N\times D}$ and $x_B \in \mathbb{R}^{M\times D}$, i.e. $x = (x_A, x_B)^T$ For the linear model, we chose a design matrix $\beta \in \mathbb{R}^{(N+M)\times 2}$ of the form

$$\beta = \begin{bmatrix} 1 & -\frac{1}{N} \\ \vdots & \vdots \\ 1 & -\frac{1}{N} \\ 1 & \frac{1}{M} \\ \vdots & \vdots \\ 1 & \frac{1}{M} \end{bmatrix} \tag{5}$$

where the first column gives the constant component and the second column (the first N entries equal to $-\frac{1}{N}$ and the second M entries equal to $\frac{1}{M}$) encodes a contrast for the difference between the datasets. The matrix is orthogonal, thus avoiding a singular design. To produce a version of the data with domain effects removed, we fit this to the data with least squares $\gamma = \min_\gamma \|A\gamma - X\|_2^2$, $\gamma \in \mathbb{R}^{2\times D}$ and project out the second component like

$$\hat{x}_{Linear} = x - x_{(:,2)}\gamma_{(2,:)} \tag{6}$$

to obtain the linearly domain-corrected data.

D.2 Harmony

For *Harmony*, we used *Harmonypy* (version 0.05) (https://github.com/slowkow/harmonypy), which is the adapted *Harmony* [25] version for the Python environment. As input, we provided a PCA

embedding of the raw data (preprocessed). Here, we used the same number of principle components (PCs) as used for RAVE. Since *Harmonypy* returns corrected PCs, we performed further evaluation on these PCs (cf. Appendix Section A.1). To find the best model(s), we performed a random search over hyperparameters. We chose the best model with Acc_{dom} close to or at chance level, while having high Acc_{type} on predicted cell type labels. Furthermore, we used the exact same dataset splits as we did for RAVE and RAVE+.

D.3 scGen

We used *scGen* [26] (version 2.0.0) within the *Scanpy* [57] (version 1.7.2) working environment. As input to *scGen*, we used the raw responses with dataset source information (either dataset *A* or *B*) using the AnnData [57] object format (version 0.7.6). To run *scGen*, we used the following functions as described in the documentation (https://scgen.readthedocs.io/en/latest/tutorials/scgen_batch_removal.html): $setup\_anndata$ to setup the AnnData object for *scGen*, $SCGEN$ to setup the model, $train$ to train the model and $batch\_removal$ to remove inter-experimental variability.

As *scGen* returns corrected input data, we performed PCA on the output data, which were used for further evaluation (cf. Appendix Section A.1). Here, we used the same number of principle components (PCs) as used for RAVE. To find the best model, we performed a random search over hyperparameters. Just like *Harmony*, we chose the best model that had Acc_{dom} close to or at chance level, while having high Acc_{type} on predicted cell type labels.

D.4 Results of Dataset-Mixing by Harmony and scGen

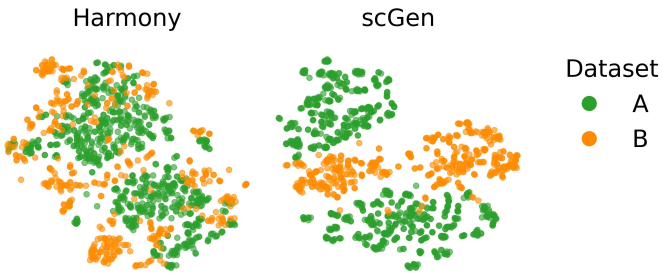

Figure 8: **Dataset Embeddings**. t-SNE embeddings of corrected data by *Harmony* (left) and *scGen* (right). Embedded cells are colored by dataset.

The low dimensional t-SNE embeddings [52, 58] (Figure 8), performed after the application of the two comparison methods (*Harmony* and *scGen*), show that cells from datasets *A* and *B* are not properly mixed; hence they are not removing inter-experimental variability sufficiently (see main paper, Table 1).

E Simulation experiments

In Figure 9, we present the results of the simulation experiments discussed in the main text. More specifically, we show example simulated cell responses for both stimuli (i.e., datasets 'A' and 'B') in Fig. 9A. Then in Fig. 9B, we demonstrate with a t-SNE embedding that the two datasets show clear inter-experimental variability. However, after correction with RAVE+, we can see in Fig. 9C that the two datasets have become aligned, and that the different cell types form clearly separated "islands". And lastly, in Fig. 9D, we see that the depth distributions of the RAVE+ corrected data are much better aligned with the ground-truth EM distributions than those of the raw data. This last steps further supports our validation procedure for RAVE+ on real data, based on EM IPL depth profiles.

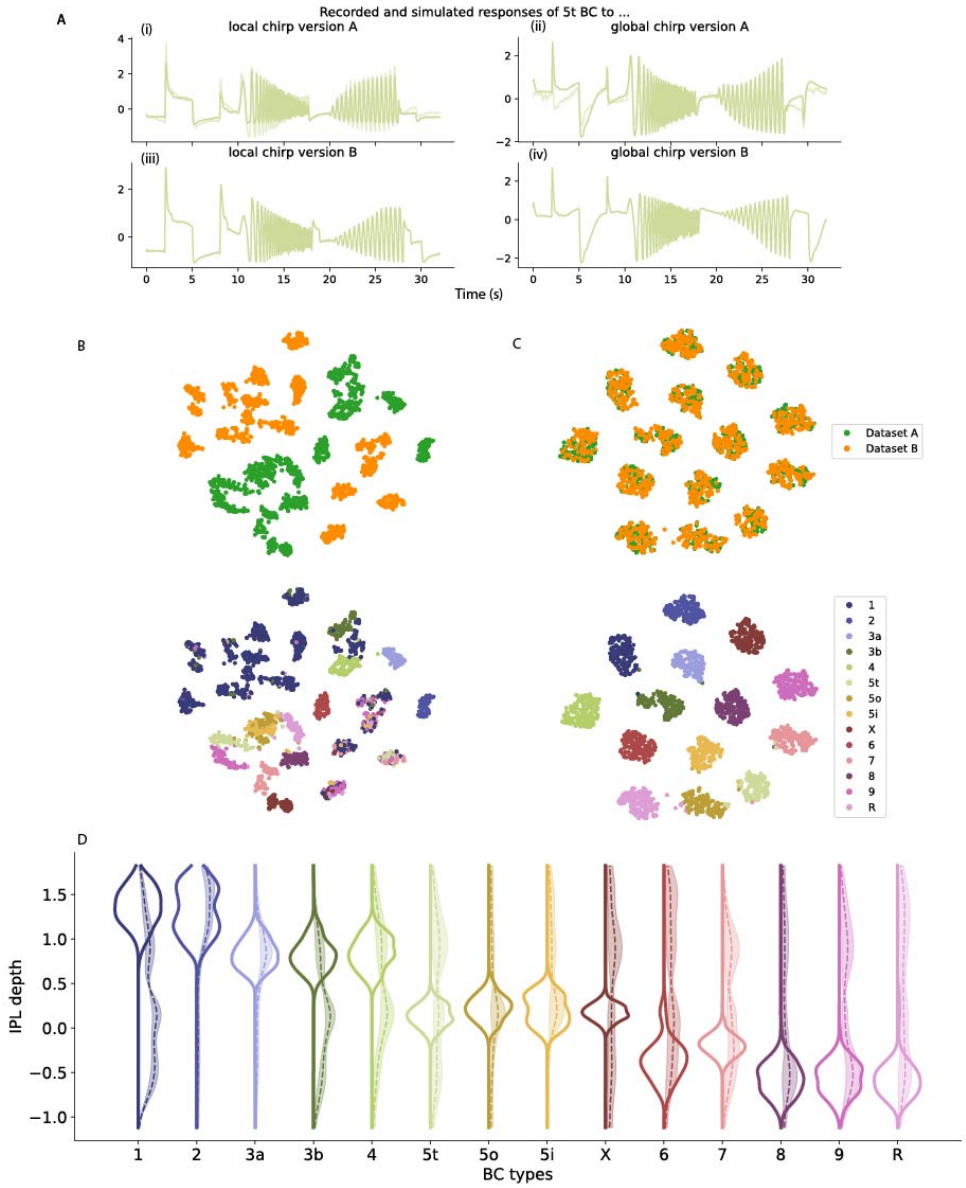

Figure 9: **RAVE+ results on simulated BC responses A**: Simulated (bold) and recorded (light) BC responses of example type 5t in response to (i) the local chirp version A (i.e. the stimulus played in dataset *A*); (ii) the global chirp version A; (iii) the local chirp version B (i.e. the stimulus played in dataset *B*); (iv) the global chirp version B. Note that for chirp version B, we do not have ground truth type labels for recorded responses. **B**: tSNE embedding of raw simulated test set data, colored according to dataset ground truth labels (top) and according to type labels predicted by a classifier trained on raw simulated responses of dataset *A* (bottom). The classifier fails for dataset *B*. **C**: same as **B**, but for `RAVE+` output. A classifier trained on `RAVE+` output for dataset A achieves accuracies of 1 for dataset *A* and 0.99 for dataset *B*. **D**: Distributions per cell type over IPL depth for EM data (distribution shown to the left), `RAVE+` output (solid line to the right) and raw data (dashed line to the right). Shaded area around the distributions shown to the right indicate SD across 10 seeds of the classifier. We sampled IPL depth values for the simulated data according to the type specific distributions known from EM data.

F Details for Performance Evaluation

F.1 Dataset-Mixing

To evaluate dataset-mixing, we used the `scikit-learn` [59] (version 0.24.1) implementation of the adjusted Rand Index (ARI) (cf. [28]).

F.2 Domain and Cell Type Classifier

In order to evaluate the model correction, we employ a domain and cell type classifier by using a random forest classifier (RFC) [60] from `scikit-learn` with cross-validated hyperparameters for each model. The RFC gets fitted on a subset of dataset *A* and validated on a held-out validation set. We performed the cross-validated grid search on the following hyperparameters: $n\_estimators$ (5, 10, 20, 30), max_depth (5, 10, 15, 20, `None`), ccp_alpha (0, 0.001, 0.01) and $max\_samples$ (0.5, 0.7, 0.9, 1). The grid search was performed using 10 random seeds to avoid overfitting (see main paper, section 4.3.1) and the best scoring RFC (highest Acc_{type}; lowest Acc_{dom} on validation set, respectively) was selected to predict cell types or domain labels on the test set of the corrected data.

F.3 Visualization of Dataset Embedding

We used the t-SNE algorithm [52] to visualize the cells in a low dimensional space [58]. For this purpose, we chose the `openTSNE` [61] implementation (version 0.6.0) in Python and ran it with default parameters and fixed seed.

Study III: Nitric oxide mediates differential effects in mouse retinal ganglion cells:

"Gonschorek et al. ‘Nitric oxide mediates differential effects in mouse retinal ganglion cells.’ (2023). in preparation."

Nitric oxide mediates differential effects in mouse retinal ganglion cells

Dominic Gonschorek[1-3], Tom Schwerd-Kleine[1-3], Jonathan Oesterle[1-2], Ryan Arlinghaus[1-2], Zhijian Zhao[2], Timm Schubert[1,2], and Thomas Euler[1-4,✉]

[1]Werner Reichardt Centre for Integrative Neuroscience, University of Tübingen, Germany
[2]Institute for Ophthalmic Research, University of Tübingen, Germany
[3]GRK 2381 'cGMP: From Benchside to Bed', University of Tübingen, Germany
[4]Bernstein Center for Computational Neuroscience, University of Tübingen, Germany

Neuromodulators have major influences on the regulation of neural circuit activity across the nervous system. Nitric oxide (NO) is a prominent neuromodulator in many circuits, and has been extensively studied in the retina across species. Here, NO has been associated with the regulation of light adaptation, gain control, and gap junction coupling, but its effect on the retinal output, specifically on the different types of retinal ganglion cells (RGCs), is still poorly understood.

Here, we used two-photon Ca^{2+} imaging to record RGC responses to visual stimuli to investigate the neuromodulatory effects of NO on the cell type-level in the *ex vivo* mouse retina. We found that about one third of the RGC types displayed highly reproducible and cell type-specific response changes during the course of an experiment, even in the absence of NO. Accounting for these adaptational changes allowed us to isolate NO effects on RGC responses. This revealed that NO affected only three RGC types, which became more active due to a reduction of their response suppression. Notably, NO had no discernible effect on their spatial receptive field size and surround strength. Together, our data suggest that NO specifically modulates suppression of the temporal response in a distinct group of contrast-suppressed RGC types. Finally, our study demonstrates the need for recording paradigms that takes adaptational, non-drug-related response changes into account when analysing potentially subtle pharmacological effects.

retina | ganglion cells | nitric oxide | neuromodulation
Correspondence: ***thomas.euler@cin.uni-tuebingen.de***

Introduction

The retina can be considered as a visual signal processor with a straightforward functional architecture (1–3). The photoreceptors convert the stream of photons from the environment into an electrical signal, which is relayed downstream by the bipolar cells (BCs) to the retinal ganglion cells (RGCs), the tissue's output neurons. Along this vertical pathway, the visual signal is shaped by two lateral, inhibitory cell classes: horizontal cells in the outer and amacrine cells (ACs) in the inner retina. The resulting intricate networks allow sophisticated computations, reflected in the >40 output channels that relay diverse visual feature representations as spike trains to higher visual brain regions (4–6).

Over the past decades, the retinal synaptic networks have been intensively studied, which greatly broadened our understanding of early visual signal processing (7–12). Specifically, the availability of connectome data for large parts of the retina helped in unprecedented ways, as can be seen from studies, for example, investigating direction-selective circuits and their precise wiring regarding starburst ACs and direction-selective RGCs (e.g., (13, 14)).

What is not well-captured by connectomic approaches are 'wireless' interactions mediated by neuromodulators, which comprise a broad variety of very different small molecules, including monoamines (e.g., dopamine (15, 16), histamine (17, 18), serotonin (19)), endocannabinoids (20–22), gasotransmitters such as nitric oxide (NO) (23, 24)), as well as a variety of neuropeptides (e.g., neuropeptide Y (25, 26)). Only few of the >20 neuromodulators (27) found in the retina are released from centrifugal fibers (e.g., histamine and serotonin (17–19, 28)), whereas most of them are released in addition to GABA or glycine by ACs (27, 29–32). Neuromodulators have long been implicated in adapting the retina to different contextual states necessary to robustly perform in a highly dynamic visual environment (e.g., (33–36)). For instance, dopamine (DA) is released by a distinct type of AC (37–40), and has been shown to regulate light adaptation (33, 41, 42) via several cellular mechanisms, such as modulation of gap-junction coupling (34, 43–45) and intrinsic conductances (46, 47). More recently, histamine was proposed to shape the retinal code in a top-down modulatory manner related to the animal's arousal state (17, 28). For several neuromodulators, the retinal release sites and receptors are known and their effects on cellular properties have been described, however, a comprehensive, function-oriented and cell-type view of these neuromodulators' functional implications and contributions to visual signal processing is only slowly emerging (e.g., (17, 48)).

One of the better-studied neuromodulators in the retina is NO. Here, neuronal nitric oxide synthesis (nNOS) is considered the dominant enzyme producing NO relevant for retinal signal processing, and – depending on species – was shown to be present in different retinal cell classes (49–57). In the mouse retina, nNOS was mainly found in specific ACs (36, 58–60). A few years ago, Jacoby et al. (36) demonstrated that one of those AC types (nNOS2-AC) controls the light-dependent release of NO in the inner retina. NO can function via two main pathways (61): (*i*) it can bind to the NO guanylate cyclase (NO-GC; also referred to as soluble guanylate cyclase (sGC)) receptor, triggering the production of the second messenger cyclic guanosine monophosphate

(cGMP), which binds to downstream targets (55, 62, 63), and (*ii*) via S-nitrosylation (cGMP-independent) by directly modifying certain receptor and channel proteins (64, 65).

The effects of NO have been primarily linked with light adaptation and the transition between scotopic and photopic signaling pathways via a number of mechanisms, including uncoupling the gap junctions between AII ACs and On-cone BCs (66), increasing the gain of On-cone BCs in response to dim light (67, 68), and modulating the gain of Off-cone BCs (69, 70). At the retina's output level, increasing NO was found to decrease On- and Off-responses in RGCs (71). Notably, genetically knocking-out nNOS also led to a reduced light sensitivity in RGCs (72). Additionally, NO has been shown to modulate RGC responses via cGMP by altering their cGMP-gated conductances (73, 74). Taken together, NO can act on different levels and via different pathways in the retina. However, the function of NO at the cell-type level and its role in early visual processing are far from understood.

Here, we systematically studied the functional role of NO and its effects on the different retinal output channels in *ex vivo* mouse retina. Surprisingly, we observed highly reproducible, cell type-specific changes in the light responses of some RGCs already in the absence of pharmacological manipulation. To account for these adaptational response changes, we developed a recording paradigm to measure 'paired' RGC responses under control and drug conditions. Here, we found that NO had a highly selective effect on a distinct set of three suppressed-by-contrast RGCs, where NO strongly and differentially reduced the suppression in their temporal responses. Yet, NO had no discernible effects on their spatial receptive field properties. Together, our data suggest that NO modulates the visual feature representation of the retinal output in a highly selective and type-specific manner.

Results

To investigate NO effects systematically on the retinal output, we performed population imaging from somata in the ganglion cell layer (GCL) of *ex vivo* mouse retina electroporated with the synthetic Ca^{2+} indicator Oregon Green BAPTA-1 (Fig. 1a-c; Methods). To identify the different types of RGCs and detect potential NO-induced response changes, we presented a set of visual stimuli, including full-field chirps, moving bars (Fig. 1d), and binary dense noise (5).

A protocol for 'paired' control/drug recordings. Previous studies have shown that retinal responses recorded with two-photon Ca^{2+} imaging can be systematically affected by experimental factors, such as excitation laser-induced activity, photoreceptor bleaching, and temporal filtering due to Ca^{2+} buffering by the fluorescent indicator (75–77). These changes can be summarized by the umbrella term 'batch effects' (a term coined in the molecular genetics field), which can confound the biological signal and potentially cause erroneous interpretations of the data (77, 78). Such batch effects may play a role when, as in our study, data are recorded in a sequential manner to infer possible drug effects.

Because we wanted to detect potentially subtle NO effects, we devised a protocol to make experiments as comparable as possible (Fig. 1e). After placing the tissue into the recording chamber, it was allowed to recover from the electroporation for 15 min, before we light-adapted the retina for 10 min by presenting the dense noise stimulus. We then selected retinal recording fields, each of which was recorded twice for the complete stimulus set in an interlaced manner (Fig. 1e). The first field was recorded twice without perturbation (*control-control pair*). For the next field, we added the drug to the perfusion medium and incubated the tissue for approx. 15 min, before recording the field for the second time (*control-drug pair*).

Using this protocol, we recorded the following paired datasets: (*i*) a control dataset to test response stability, i.e., Ctrl 1 and Ctrl 2, (*ii*) a strychnine dataset to test the protocol for a drug with well-described effects, i.e., Ctrl 1 and Strychnine (1 µM), and *(iii)* a NO dataset to infer NO-induced response changes, i.e., Ctrl 1 and NO (DETA/NO; 100 µM). The control dataset was leveraged to reveal NO-induced effects on the background of potential nonspecific response changes over the course of the experiment.

For the following analyses, we used 3,864 paired RGCs (n_{Ctrl}=1,590; n_{NO}=1,838, $n_{strychnine}$=436) that fulfilled our response quality filtering (see Methods).

Identifying functional RGC types using a classifier. The mouse retina contains more than 40 RGC types (5, 6). As we wanted to investigate if the tested drugs differentially affect the different retinal output channels, we applied an RGC classifier (Fig. 2a) (79), which had been trained and validated on a previously published RGC Ca^{2+} imaging dataset (5). The classifier predicts an GCL cell's functional type based on soma size and the responses to chirp and moving bar stimuli (Methods). While the classifier also identifies displaced ACs, we focused our analysis on RGCs. To match the conditions under which the classifier's training data was acquired as closely as possible, we predicted types based on the responses from the first control recording (Ctrl 1). To minimize classification uncertainty, we additionally discarded cells with low confidence scores (< 0.25, see Methods for details).

We found that the distributions of the predicted RGC types in our datasets matched that of the earlier dataset quite well (Fig. 2b-d). Also the predicted mean traces for the individual RGC types were very similar to those in Baden et al. (5) (Fig. 2e), as indicated by the high correlations of their chirp and moving bar responses (Fig. 2f left and right, respectively). That the moving bar responses were more strongly correlated than the chirp responses is likely due to the lower complexity and shorter duration of the former stimulus. Nonetheless, we found the RGC classification overall to be robust.

Testing and evaluating the recording protocol. To test our recording protocol, we used the glycine receptor antagonist strychnine (Fig. S1). In the mouse retina, glycine is

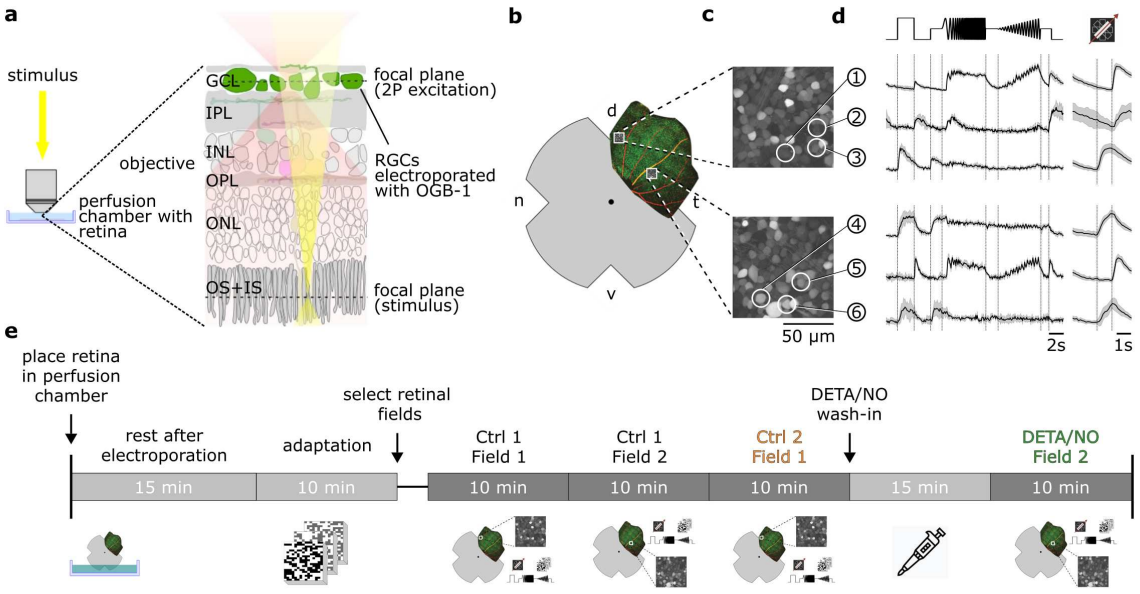

Fig. 1. Overview of experimental setup and recording procedure. (**a**) Two-photon imaging of ganglion cell layer (GCL) somata in the whole-mounted mouse retina. (**b**) Schematic *ex vivo* whole-mounted retina with two representative recording fields (white boxes; dot marks optic disc; d, dorsal; t, temporal; v, ventral; n, nasal). (**c**) Recording fields from (b) showing GCL somata loaded with Ca^{2+} indicator OGB-1 (Methods). (**d**) Representative GCL cell Ca^{2+} activity (white circles in (c)) in response to chirp (left) and moving bar stimulus (right) (black, mean; gray, s.d.). (**e**) Timeline of experimental procedure; for details, see text. IPL, inner plexiform layer; INL, inner nuclear layer; OPL, outer plexiform layer; ONL, outer nuclear layer; OS+IS, outer and inner segments; DETA/NO, nitric oxide (NO) donor.

released by small-field ACs and relays inhibition vertically across the inner plexiform layer (IPL; (32, 80)). For instance, such cross-over inhibition mediated by AII ACs relays the signals from rod BCs to Off-cone BCs, which is why strychnine eliminates Off RGC responses in dark-adapted retina (81). In the light-adapted retina, On-cone BCs boost light-Off responses in Off-cone BCs through cross-over inhibition (82, 83) and hence, strychnine affects Off-response components in RGCs – in line with our observations (Fig. S1): Strychnine application revealed additional On-responses in Off (e.g., G_1, G_2) and On-Off RGCs (e.g., G_{12}), as can be seen, for instance, in their leading-edge response to the moving bar (Fig. S1; left panel, middle column). In On RGCs, we did not detect substantial response changes with strychnine (Fig. S1; e.g., G_{17}).

Together, the strychnine dataset showcases that we can resolve drug-related effects on light responses at cell-type level.

Certain RGC types display adaptational response changes. To test if our recording conditions were stable and to exclude major batch effects, we first compared the responses of the control datasets (Ctrl 1 vs. Ctrl 2). To this end, we computed the difference between the Ctrl 1 and Ctrl 2 mean responses ($\Delta R_{(Ctrl2-Ctrl1)}$) to chirp and moving bar stimuli for each cell of every RGC type. This allowed us to quantify if and how the responses changed over the time-course of an experiment (cf. protocol in Fig. 1e). Here, we only considered RGC types with >10 paired cells (21/32).

Surprisingly, while the majority of RGC types featured stable responses (e.g., G_1, G_{21}; Fig. 3a), a substantial number of RGC types (∼34%) changed their responses to chirp and/or moving bar stimuli in the absence of any pharmacological perturbation in a highly reproducible manner (Fig. 3b). For instance, for Ctrl 2, G_{23} showed reduced responses, whereas G_{31} showed an increased response activity. Interestingly, cells assigned to the functional groups of 'Off' and 'Fast-On' RGC types displayed exclusively stable responses, whereas 'On-Off', 'Slow-On', and 'Uncertain RGCs' included types with changing responses (50%, 50%, and 67%, respectively). This diversity argues against a systematic effect (such as, e.g., general run-down) and for a cell type-specific phenomenon, which in the following we refer to as 'adaptational response changes'.

NO affects retinal output in a highly type-specific manner. Next, we investigated the effects of NO on the RGC responses. As with the control dataset, we computed the cell-wise response differences between Ctrl 1 and NO responses ($\Delta R_{(NO-Ctrl1)}$). Similar to the control dataset, the majority of RGC types displayed stable responses (e.g., G_2, G_{17}; Fig. 3c), while a smaller fraction (∼43%) changed their responses significantly (e.g., G_{28}, G_{32}; Fig. 3d) following the NO perfusion. We found that the percentage of changing types per functional group was similar to that in the control dataset: 'Off' (0%), 'Fast On' (34%), 'On-Off' (50%), 'Slow On' (66%), and 'Uncertain RGCs' (66%). This raised the question if the observed changes in the NO dataset indeed reflected NO-induced modulations or mostly adaptational response changes (as observed in the control dataset). We therefore tested for each RGC type if the response changes observed for control ($\Delta R_{(Ctrl2-Ctrl1)}$) and NO

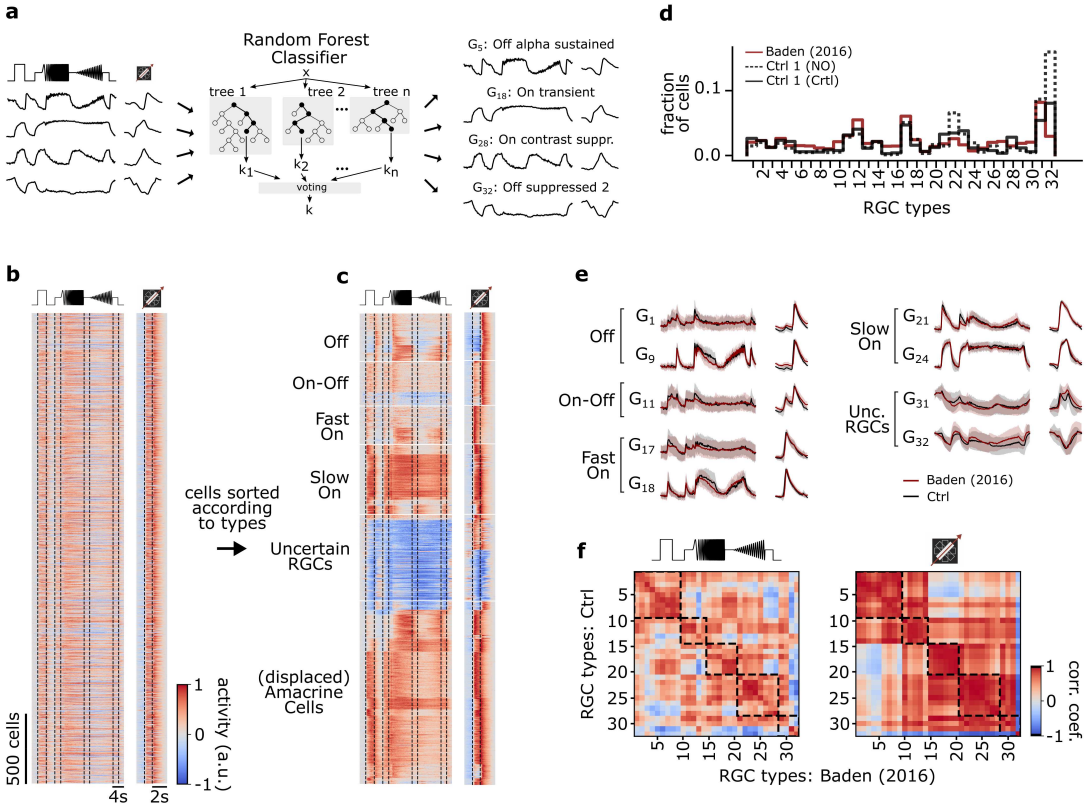

Fig. 2. Functional classification of mouse RGC types. (**a**) Illustration of the random forest classifier (RFC) to predict cell type labels for Ctrl 1 of both datasets. For each cell, Ca^{2+} responses to chirp and moving bar, soma size, and p-value of permutation test for direction selectivity (left) constitute the input to the RFC (centre) to predict a cell type label, i.e., a type G_X (right). For details, see Methods and (79). (**b**) Pooled heat map of unsorted cell responses to chirp and moving bar stimulus from both Ctrl 1 datasets after quality filtering (QI_{MB}>0.6 or QI_{chirp}>0.45, and classifier confidence score ≤ 0.25). Color bar indicates normalized response activity. (**c**) Heat map from (b), but sorted according to their assigned type. (**d**) Distribution of RGC types predicted by the RFC classifier for both Ctrl 1 of control pair (Ctrl 1 (Ctrl); solid black), of DETA/NO pair (Ctrl 1 (NO); dotted black), and for the training dataset from Baden et al. (5) (red). (**e**) Representative RGC type mean responses to chirp and moving bar (Ctrl, black; training dataset, red). (**f**) Correlation matrix of type mean responses per RGC type between Ctrl and training dataset for chirp (left) and moving bar (right). Dashed boxes indicate functional groups (Off, On-Off, Fast On, Slow On, and Uncertain RGCs; see (5)). Color bar indicates correlation coefficient.

($\Delta R_{(NO-Ctrl1)}$) were significantly different (Fig. 3e). To our surprise, this was only the case for G_{32} ('Off-suppressed 2') RGCs, suggesting highly type-selective NO effects – at least for temporal responses to chirp and moving bar stimuli.

Next, we leveraged the control dataset to disentangle NO-induced from adaptational effects at the level of the response features. To this end, we subdivided the chirp stimulus into 6 features ((i) On, (ii) Off, (iii) low frequency, (iv) high frequency, (v) low contrast, (vi) high contrast), and the moving bar into 2 features ((vii) On, (viii) Off) (Fig. 4c). Then, for every cell type and every stimulus feature, we computed the difference of the mean responses between the first and second recording separately for the control (Fig. 4a; left panel) and NO dataset (Fig. 4a; middle panel). To isolate NO-induced effects, we computed the differences ($\Delta R_{(NO-Ctrl1)} - \Delta R_{(Ctrl2-Ctrl1)}$; Fig. 4a; right panel), based on the assumption that the adaptational component of the changes would be similar for both datasets. Through this analysis, we could distinguish three response behaviors across the RGC types: (1) not NO-affected/stable, (2) showing cell type-specific adaptation, and (3) NO-affected (Fig. 4b).

As before, we found that only G_{32} is strongly affected by NO; it showed no adaptational response changes, yet its activity increased during NO application (Fig. 4b; (3)). Interestingly, a cell type that is also suppressed-by-contrast, G_{31} ('Off-suppressed 1'), displayed a similar behavior during the NO application (Fig. 4b; (2); bottom). However, G_{31} showed this response change already in the control, suggesting that this effect was adaptational and not NO-related (Fig. 4a; right panel). Notably, RGC types that were assigned to the group of the so-called 'Slow On' types (G_{21}-G_{28}), which exhibit strong and sustained responses during the frequency and contrast sequences of the chirp stimulus, showed a decrease in activity in both datasets (e.g., G_{24}; Fig. 4b; (2); top). Consequently, the changes in these response features (iii-vi) are likely adaptational (Fig. 4a) – as the changes can be found in the control (Fig. 4a; left panel) as well as in the NO dataset (Fig. 4a; middle panel), but not in the difference (Fig. 4a; right panel). Our conclusion that both adaptational and NO-mediated effects on RGC responses are highly cell type-selective is further supported by the fact that we also

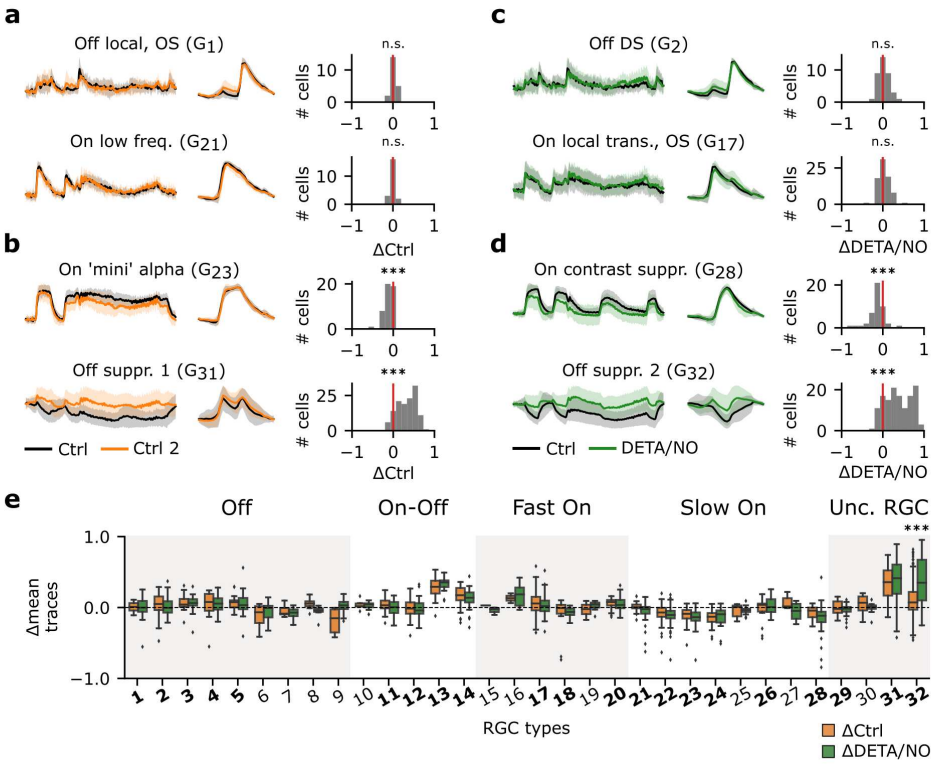

Fig. 3. Certain RGC types are affected by adaptational and/or NO-induced effects, others are unaffected. (**a**) Left: Two representative mean calcium responses of paired RGC types showing no differences between Ctrl 1 (black) and Ctrl 2 (orange) (top: G_1; bottom: G_{21}). Right: Corresponding histograms displaying the trace differences between paired cell responses of the respective cell types. Zero indicates no difference between the response of the same cell across both recordings, whereas negative values indicate a decreased and positive ones an increased activity. (**b**) Two representative RGC types that show decreased (top: G_{23}) and increased (bottom: G_{31}) response activity during Ctrl 2. (**c**) As in (a), but between paired Ctrl 1 (black) and DETA/NO (green) (top: G_2; bottom: G_{17}). (**d**) As (c), but showing two cell types that display a decreased (top: G_{28}) and increased (bottom: G_{32}) activity when perfused with DETA/NO. ***: p<0.001; One-sample T-test. (**e**) Box plots of trace differences of all paired cells of all RGC types from control (Ctrl 1 & Ctrl 2; orange) and NO dataset (Ctrl & DETA/NO; green). Bold numbers indicate RGC types with >10 paired cells per dataset and condition. Dashed line shows zero baseline, i.e., no difference between traces. Gray and white background blocks summarize the larger functional groups for better visualization (Off, On-Off, Fast On, Slow On, Uncertain RGCs). ***: p<0.001; Mann-Whitney U-Test.

found several RGC types that showed stable responses during control and NO application (Fig. 4b; (1)).

Clustering of G_{32} responses reveals three functionally distinct RGC types with different NO-sensitivity. According to Baden et al. (5), G_{32} features a coverage factor of ∼4. As the average coverage factor of mouse RGCs was estimated to be ∼2 (5, 84), G_{32} likely consists of several RGC types. To test this, we performed Mixture of Gaussian clustering of the RGC cells assigned as G_{32} (Fig. 5a-c) using the Ctrl 1 responses to chirp and moving bar stimuli from both datasets (Fig. 5a). We found three distinct clusters – all suppressed-by-contrast but to different degrees (Fig. 5d,e).

All three G_{32} clusters showed little to no adaptation for the control dataset, but displayed differential modulations in response to NO application, with cluster 1 exhibiting the strongest NO effect (Fig. 5f,g; left) for both stimuli. The NO effect was statistically significant in clusters 1 and 2 – both for mean trace difference and correlation (Fig. 5g; left and right, respectively) – but not in cluster 3.

Since the prominent feature of these RGC types is suppression by stimulus contrast, we compared their suppression strength between the conditions using a suppression index (SI; see Methods). Here, we found that in cluster 1, SI did not change between Ctrl 1 and Ctrl 2, but was significantly reduced with NO. In fact, with NO, these cells lost their suppressive feature and were rather excited than suppressed by the stimuli. Cluster 2 displayed significant differences in SI in both control and NO dataset.

Taken together, our data indicate that G_{32} may consist of three suppressed-by-contrast RGC types and that in at least two of them, the contrast suppression is strongly modulated by NO. This subdivision of G_{32} into three types is supported by its high coverage factor (5, 84).

NO does not affect RGC receptive field properties. So far, we focused on effects in the temporal response domain, where we found that mainly G_{32} types were affected by an increase in NO levels. However, NO has been shown to affect electrical coupling, e.g., by reducing conductance between AII ACs and On-cone BCs (66), and hence may alter receptive field (RF) properties. Therefore, we next investigated the effects of NO on the spatial RFs (sRFs) of the individ-

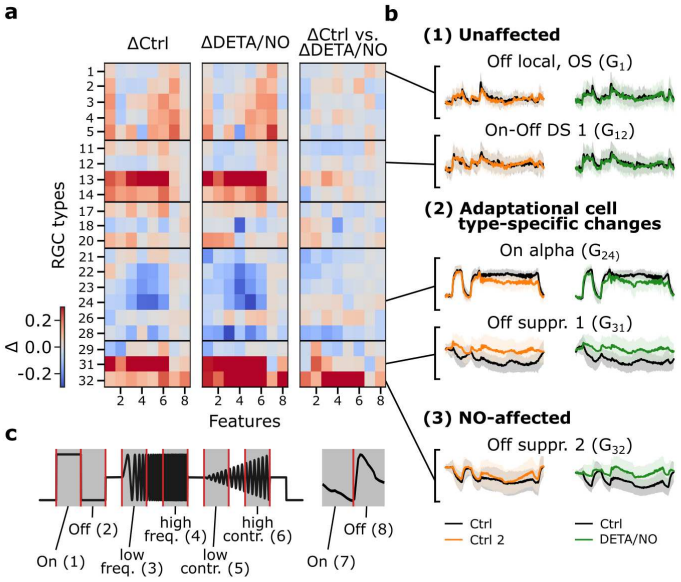

Fig. 4. Disentangling NO-induced effects from adaptational response changes reveals type-specific NO modulation. (**a**) Left: Difference between paired Ctrl 2 and Ctrl 1 RGC traces per type subdivided into 8 features ($\Delta R_{(Ctrl2-Ctrl1)}$). Color-code indicates response increase (red), no change (white) and decrease (blue) for Ctrl 2. Middle: Difference between DETA/NO and Ctrl ($\Delta R_{(NO-Ctrl1)}$). Right: Difference between the two heatmaps ($\Delta R_{(NO-Ctrl1)} - \Delta R_{(Ctrl2-Ctrl1)}$). (**b**) Example chirp traces categorized into unaffected (top two types: G_1, G_{12}), adaptational (two middle types: G_{24}, G_{31}) and NO-affected (bottom: G_{32}). Left traces show exemplary responses per type from the control dataset (black: Ctrl 1; orange: Ctrl 2) and NO dataset (black: Ctrl; green: DETA/NO). (**c**) Subdividing the chirp (left) and moving bar (right) stimuli into 8 features for detailed feature analysis. The chirp is subdivided into 6 features ((i) on, (ii) off, (iii) low frequency, (iv) high frequency, (v) low contrast and (vi) high contrast); the moving bar into 2 ((vii) on and (viii) off).

ual RGC types. To this end, following the same experimental paradigm as described earlier, we recorded RGC responses to binary dense noise. Next, we computed their sRFs for both recording conditions (Fig. 6a) using spike-triggered averaging (85), obtaining control and NO sRFs, and then fitted a Gaussian to each sRF's center (Fig. 6a). We focused the following analysis on RGC types with reliable sRF estimates (see Methods). These included types in all five larger RGC groups (5): G_5 for Off; G_{11}, G_{12}, and G_{14} for On-Off; G_{17} for Fast On; G_{23}, G_{24}, G_{26}, and G_{27} for Slow On; G_{31} and G_{32} for Uncertain RGCs. In these types, sRFs were very stable in both control and NO condition.

Using the difference in sRF center diameter between control and NO as a metric (Fig. 6b), we did not find NO to cause any significant changes in sRF size in any of the analysed RGC types. Next, we tested if the sRF surround was affected by NO, because a modulation of inhibitory synaptic input and/or electrical coupling may cause a change in surround strength. As a measure of surround strength, we computed a surround index (see Methods) of control and NO sRFs (Fig. 6c). Like with the sRF center diameter, we did not find significant differences for any of the analysed RGC types. Surprisingly, also G_{32} did not show NO-mediated differences in its sRF properties, implying that NO may only affect its temporal response, but not its RF organization.

Taken together, at least for the tested RGC types, we did not detect any significant NO effects neither on sRF center size not surround strength.

Discussion

We used two-photon Ca^{2+} imaging to record RGC responses to various visual stimuli to investigate the neuromodulatory effects of elevated NO levels on the signal processing across RGC types in the mouse retina. To our surprise, already without pharmacological perturbation, we found that about one third of the RGC types displayed highly reproducible and cell type-specific response changes during the course of an experiment – a finding of potentially high relevance especially for pharmacological experiments in the *ex vivo* retina. Accounting for these adaptational changes enabled us to isolate NO-related effects on RGC responses. Here, we revealed that mainly the RGCs assigned to G_{32} (Off suppressed 2) were affected by NO, which strongly reduced the response suppression and rendered the cells more active. Further, we demonstrated that G_{32} likely consists of three types – consistent with its high coverage factor (5) – that were all differentially modulated by NO. Finally, for a representative subset of RGC types, we showed that elevating NO levels has no discernible effect on sRF size or surround strength. Together, our data suggest that NO specifically modulates response suppression in a group of contrast-suppressed RGC types. Additionally, our study demonstrates the need for recording paradigms that takes adaptational, non-drug-related response changes into account when analysing potentially subtle pharmacological effects.

Nitric oxide as a modulator in the retina. Neuronal nitric oxide synthesis (nNOS) has been detected in different retinal cell classes (49–57). In the mouse retina, light-dependent

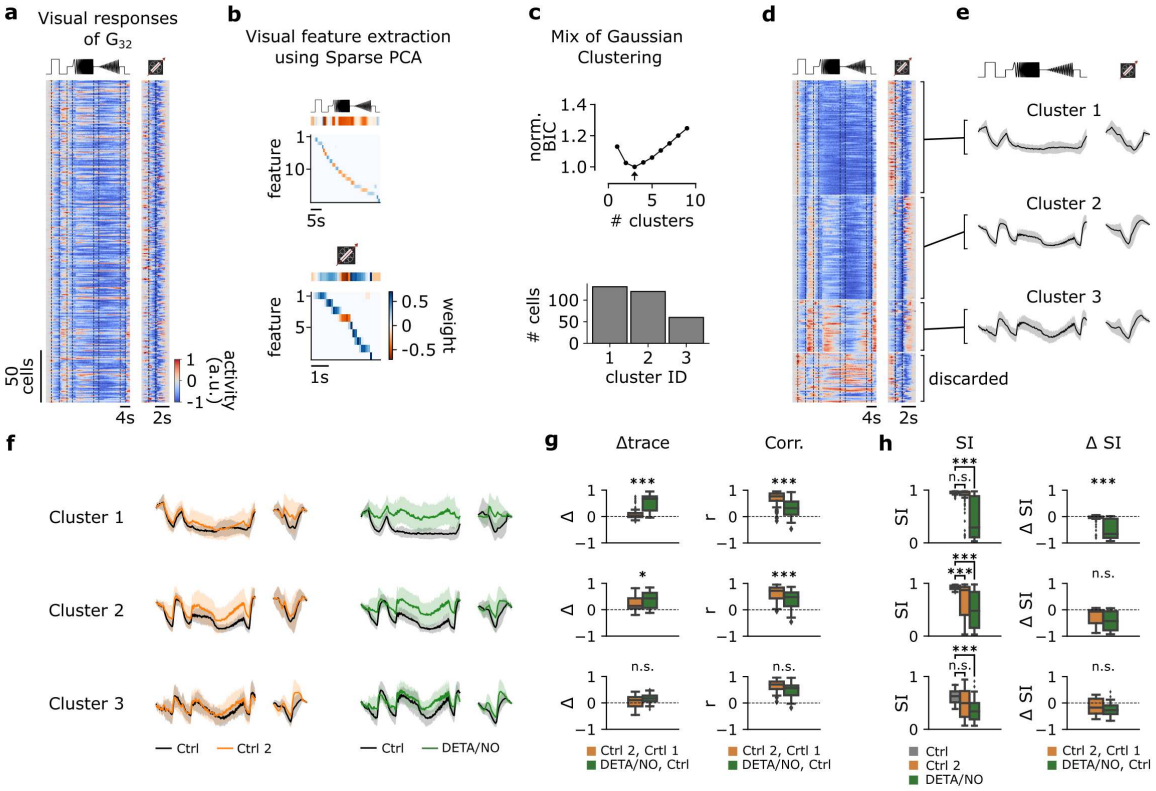

Fig. 5. Functional clustering of the G_{32} reveals three distinct types that are differently affected by NO. (**a**) Visual responses of G_{32} cells recorded from several experiments in response to the full-field chirp (left) and moving bar (right) stimuli. (**b**) Visual features extracted from chirp (top) and moving bar (bottom) stimuli using sparse PCA on the responses. Color indicates weight of each feature. (**c**) Top: Bayesian Information Criterion (BIC) as function of number of clusters. Arrow indicates lowest BIC and the number of clusters to choose. Bottom: Number of cells per predicted cluster. (**d**) Cells sorted according to their assigned cluster. Cells at the bottom were discarded. (**e**) Mean responses of the 3 corresponding clusters for the chirp (left) and moving bar (right). (**f**) Left: Paired mean responses of the 3 clusters to Ctrl 1 (black) and Ctrl 2 (orange). Right: Cluster mean responses to Ctrl 1 (black) and DETA/NO (green) (**g**) Left: Trace difference between Ctrl 2 & Ctrl 1 (orange) and DETA/NO & Ctrl (green) for the 3 clusters (cluster 1-3 from top to bottom). Right: Correlation coefficient between Ctrl 2 & Ctrl 1 (orange) and DETA/NO & Ctrl 1 (green) for the 3 clusters. ***: p<0.001; independent T-test & Mann-Whitney U-Test. (**h**) Left: Suppression index (SI) computed for Ctrl 1 (gray), Ctrl 2 (orange) and DETA/NO (green) for the 3 clusters. ***: p<0.001; Kruskal-Wallis test & Dunnett's test. Right: Difference of SI between Ctrl 2 & Ctrl 1 (orange) and DETA/NO & Ctrl 1 (green). ***: p<0.001; independent T-test & Mann-Whitney U-Test.

NO production seems to mainly occur in specific AC types (36). The main NO-sensor (NO-GC), which connects NO to intracellular cGMP signalling (55, 60, 86, 87) is present in all retinal layers (70). This and earlier findings of NO modulating response gain in BCs (68, 69) and RGCs (71), led us to expect global NO effects on retinal output. What we found instead was a highly selective effect in a distinct group of likely three types of RGCs formerly amalgamated in G_{32} (5). Specifically, NO affected these types in a similar way, yet differentially by increasing their activity and partly eliminating their characteristic feature of being suppressed by temporal contrast. Notably, the NO effect was only visible in their temporal responses; we did not detect changes in their sRF properties. This RGC type-selectivity of NO neuromodulation is reminiscent of a recent study, where dopamine was found to modulate distinct response features of specific RGC types (48).

To ensure reliable RGC type classification, our set of visual stimuli was restricted to artificial ones (full-field chirps, moving bards, and dense noise). Such artificial stimuli probe the stimulus space in a rather selective and limited fashion. Hence, we cannot exclude that we missed NO effects that may have come apparent for other, more complex stimuli. Specifically, natural images or movies, stimuli that are closer to what the retina evolved to process (88), may be needed for a more complete picture of the functional implication of retinal neuromodulation. That natural stimuli can reveal novel nonlinear properties of retinal functions was demonstrated, for example, by Goldin et al. (89), who showed that a RGC's contrast selectivity can be context-dependent. Similarly, Höfling et al. (90) discovered a RGC type that may serve as a 'horizon detector'. This finding was made possible by combining natural movie-evoked RGC responses and a convolutional neural network trained on these data. These approaches highlight that future studies on neuromodulators should also employ natural stimuli.

Another important aspect to consider when studying NO neuromodulation in the retina is the level of light adaptation. Several studies proposed that NO facilitates the transition across light levels (66, 68, 69), especially to photopic con-

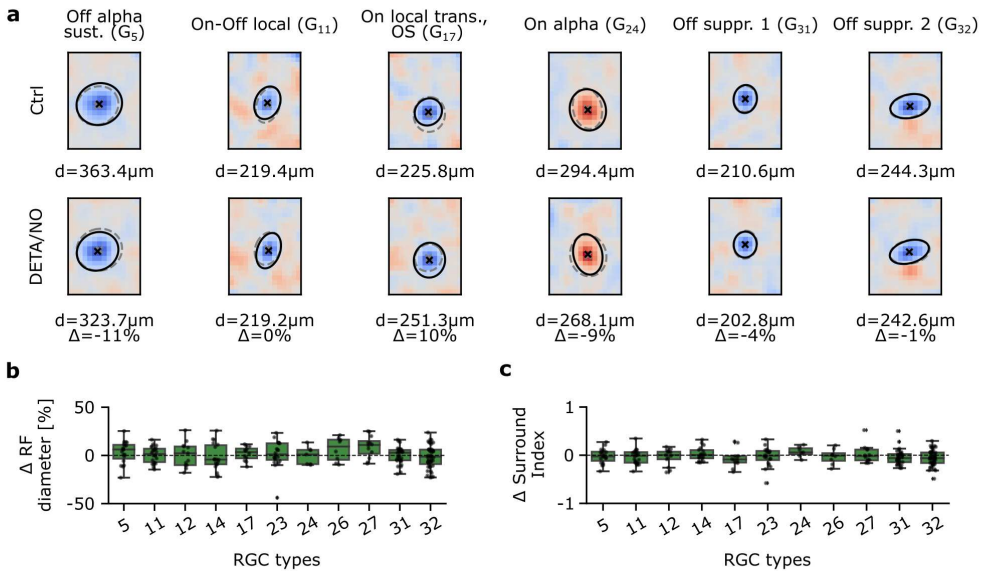

Fig. 6. Spatial receptive fields are not affected by NO across various RGC types. (**a**) Representative estimated spatial receptive fields (sRF) of six RGC types. Top: Estimated sRFs to Ctrl 1. Cross indicates RF center; solid line indicates outline of the Gaussian fit of the RF center; dashed outline indicates corresponding Gaussian fit of the same cell to DETA/NO. Bottom: Same as top, but of DETA/NO condition. (**b**) RF diameter difference in percentage between DETA/NO and Ctrl 1. Only types with more than 5 paired cells were included. One-sample T-test. (**c**) Surround index difference between DETA/NO and Ctrl 1. Only types with more than 5 paired cells were included. One-sample T-test.

ditions (36). Since we employed two-photon imaging, which inevitably results in a certain level of background 'illumination' (see discussion in (75, 91)), our experiments were performed in the low photopic range. Therefore, it is possible that NO-mediated neuromodulation serves different, light-level dependent functions: more globally during the transition from scotopic to mesopic/photopic, and more cell type-specific in the photopic regime – as we reported here.

Adaptational, cell type-specific response changes. Every recording method introduces technique-specific biases that have to be considered in the data analysis and interpretation. For two-photon imaging with fluorescent Ca^{2+} sensors, these potential biases include Ca^{2+} buffering, sensor bleaching by the excitation laser and, in the case of bulk-loading with synthetic dyes (as in our experiments; see also (92)), slow leakage of indicator from the cells. In retinal imaging, additional potentially confounding factors are an excitation laser-induced baseline activity and photoreceptor bleaching (76, 91). These biases are expected to be systematic, e.g., causing a decrease in signal-to-noise ratio across (RGC) responses (i.e., run-down). To account for this, we established a recording paradigm that produced a control and a NO dataset consisting of paired RGC responses.

When analyzing the control dataset, we were surprised by finding response changes in approx. a third of the RGC types already in the absence of the NO donor. These changes were consistent for a particular type but differed between types. Typically, we did not observe a simple overall decrease or increase in activity, but rather selective changes of response features: For instance, in G_{24} (Slow On) only its response to high frequency and low contrast was reduced, while the remaining response features were not affected. Together, this strongly argues against a systematic, recording technique-related bias. Currently, we can only speculate about the mechanism(s) underlying this type-specific adaptation. The most parsimonious explanation may be related to the *ex vivo* condition of the retina: While we allowed the tissue to settle and adapt to perfusion medium, light level, temperature etc. for approx. 25 min, extracellular signaling molecules, such as neuromodulators, may be depleted and washed out over the course of the experiment, resulting in differential adaptation of various RGC types. In any case, as type-selective adaptations can confound the recorded responses in a complex manner, a paired recording paradigm as the one described here, is recommended – in particular for pharmacological experiments.

Combining population imaging, RGC classification and paired recordings to study neuromodulation of retinal output. In this study, we investigated the neuromodulatory effects of NO on the retinal output signal. To this end, we combined experimental and computational approaches to dissect NO-mediated effects at the RGC type-level. The latter is important for understanding neuromodulator function for early vision, because the visual information is sent to the

brain via parallel feature channels, represented by >40 RGC types in the mouse retina (5, 6, 84, 93). We demonstrated that our approach enable adaptational from actual NO-induced effects using a rather simple and straightforward linear analysis (i.e., focusing on mean trace or RF size differences), which means that we may have missed potential nonlinear effects. However, the here presented pipeline for analyzing neuromodulation in neural circuits constitutes a framework that can be easily extended by more advanced analyses.

ACKNOWLEDGEMENTS
We thank G. Eske & M. Harrer for technical support, and O. Marre, M. Goldin, C. Makino & R. Feil for fruitful discussions. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) 335549539/GRK2381.

Bibliography

1. Heinz Wassle and Brian B Boycott. Functional architecture of the mammalian retina. *Physiological reviews*, 71(2):447–480, 1991.
2. Richard H Masland. The fundamental plan of the retina. *Nature neuroscience*, 4(9):877–886, 2001.
3. Heinz Wässle. Parallel processing in the mammalian retina. *Nature Reviews Neuroscience*, 5(10):747–757, 2004.
4. Tania A Seabrook, Timothy J Burbridge, Michael C Crair, and Andrew D Huberman. Architecture, function, and assembly of the mouse visual system. *Annual review of neuroscience*, 40:499–538, 2017.
5. Tom Baden, Philipp Berens, Katrin Franke, Miroslav Román Rosón, Matthias Bethge, and Thomas Euler. The functional diversity of retinal ganglion cells in the mouse. *Nature*, 529: 345–350, 2016. ISSN 9781137332875. doi: 10.1038/nature16468.
6. J. Goetz, Z. F. Jessen, A. Jacobi, A. Mani, S. Cooler, D. Greer, S. Kadri, J. Segal, K. Shekhar, J. R. Sanes, and G. W. Schwartz. Unified classification of mouse retinal ganglion cells using function, morphology, and gene expression. *Cell Rep*, 40(2):111040, 2022. ISSN 2211-1247 (Electronic). doi: 10.1016/j.celrep.2022.111040. Goetz, Jillian Jessen, Zachary F Jacobi, Anne Mani, Adam Cooler, Sam Greer, Devon Kadri, Sabah Segal, Jeremy Shekhar, Karthik Sanes, Joshua R Schwartz, Gregory W eng 2022/07/14 Cell Rep. 2022 Jul 12;40(2):111040. doi: 10.1016/j.celrep.2022.111040.
7. James R Anderson, Bryan W Jones, Carl B Watt, Margaret V Shaw, Jia-Hui Yang, David DeMill, James S Lauritzen, Yanhua Lin, Kevin D Rapp, David Mastronarde, et al. Exploring the retinal connectome. *Molecular vision*, 17:355, 2011.
8. Robert E Marc, Bryan W Jones, Carl B Watt, James R Anderson, Crystal Sigulinsky, and Scott Lauritzen. Retinal connectomics: towards complete, accurate networks. *Progress in retinal and eye research*, 37:141–162, 2013.
9. Moritz Helmstaedter, Kevin L Briggman, Srinivas C Turaga, Viren Jain, H Sebastian Seung, and Winfried Denk. Connectomic reconstruction of the inner plexiform layer in the mouse retina. *Nature*, 500(7461):168–174, 2013.
10. Moritz Helmstaedter. Cellular-resolution connectomics: challenges of dense neural circuit reconstruction. *Nature methods*, 10(6):501–507, 2013.
11. Felice A Dunn and Rachel OL Wong. Wiring patterns in the mouse retina: collecting evidence across the connectome, physiology and light microscopy. *The Journal of physiology*, 592(22):4809–4823, 2014.
12. Jamie Johnston and Leon Lagnado. General features of the retinal connectome determine the computation of motion anticipation. *Elife*, 4:e06250, 2015.
13. Kevin L Briggman, Moritz Helmstaedter, and Winfried Denk. Wiring specificity in the direction-selectivity circuit of the retina. *Nature*, 471(7337):183–188, 2011.
14. Huayu Ding, Robert G Smith, Alon Poleg-Polsky, Jeffrey S Diamond, and Kevin L Briggman. Species-specific wiring for direction selectivity in the mammalian retina. *Nature*, 535(7610):105–110, 2016.
15. Paul Witkovsky. Dopamine and retinal function. *Documenta ophthalmologica*, 108:17–39, 2004.
16. Suva Roy and Greg D Field. Dopaminergic modulation of retinal processing from starlight to sunlight. *Journal of pharmacological sciences*, 140(1):86–93, 2019.
17. Rebekah A Warwick, Serena Riccitelli, Alina S Heukamp, Hadar Yaakov, Lea Ankri, Jonathan Mayzel, Noa Gilead, Reut Parness-Yossifon, and Michal Rivlin-Etzion. Top-down modulation of the retinal code via histaminergic neurons of the hypothalamus. *bioRxiv*, pages 2022–04, 2022.
18. Ursula Greferath, M Kambourakis, Christian Barth, EL Fletcher, and M Murphy. Characterization of histamine projections and their potential cellular targets in the mouse retina. *Neuroscience*, 158(2):932–944, 2009.
19. Justine Masson. Serotonin in retina. *Biochimie*, 161:51–55, 2019.
20. Stephen Yazulla. Endocannabinoids in the retina: from marijuana to neuroprotection. *Progress in retinal and eye research*, 27(5):501–526, 2008.
21. Thomas Schwitzer, Raymund Schwan, Karine Angioi-Duprez, Anne Giersch, and Vincent Laprevote. The endocannabinoid system in the retina: from physiology to practical and therapeutic applications. *Neural plasticity*, 2016, 2016.
22. Charles F Yates, Jin Y Huang, and Dario A Protti. Tonic endocannabinoid levels modulate retinal signaling. *International Journal of Environmental Research and Public Health*, 19 (19):12460, 2022.
23. Ira M Goldstein, Philipp Ostwald, and Steven Roth. Nitric oxide: a review of its role in retinal function and disease. *Vision research*, 36(18):2979–2994, 1996.
24. Alex H Vielma, Mauricio A Retamal, and Oliver Schmachtenberg. Nitric oxide signaling in the retina: what have we learned in two decades? *Brain research*, 1430:112–125, 2012.
25. Ana Santos-Carvalho, Ana Rita Alvaro, Joao Martins, Antonio Francisco Ambrosio, and Claudia Cavadas. Emerging novel roles of neuropeptide y in the retina: from neuromodulation to neuroprotection. *Progress in neurobiology*, 112:70–79, 2014.
26. Ana Santos-Carvalho, António Francisco Ambrósio, and Cláudia Cavadas. Neuropeptide y system in the retina: From localization to function. *Progress in retinal and eye research*, 47:19–37, 2015.
27. Wenjun Yan, Mallory A Laboulaye, Nicholas M Tran, Irene E Whitney, Inbal Benhar, and Joshua R Sanes. Mouse retinal cell atlas: molecular identification of over sixty amacrine cell types. *Journal of Neuroscience*, 40(27):5177–5195, 2020.
28. Matthew J Gastinger, Ning Tian, Tamas Horvath, and David W Marshak. Retinopetal axons in mammals: emphasis on histamine and serotonin. *Current eye research*, 31(7-8): 655–667, 2006.
29. DONALD M O'Malley and Richard H Masland. Co-release of acetylcholine and gamma-aminobutyric acid by a retinal neuron. *Proceedings of the National Academy of Sciences*, 86(9):3414–3418, 1989.
30. Donald M O'Malley, Julie H Sandell, and Richard H Masland. Co-release of acetylcholine and gaba by the starburst amacrine cells. *Journal of Neuroscience*, 12(4):1394–1408, 1992.
31. Hajime Hirasawa, Rebecca A Betensky, and Elio Raviola. Corelease of dopamine and gaba by a retinal dopaminergic neuron. *Journal of Neuroscience*, 32(38):13281–13291, 2012.
32. Jeffrey S Diamond. Inhibitory interneurons in the retina: types, circuitry, and function. *Annual review of vision science*, 3:1–24, 2017.
33. Michael D Flood, Johnnie M Moore-Dotson, and Erika D Eggers. Dopamine d1 receptor activation contributes to light-adapted changes in retinal inhibition to rod bipolar cells. *Journal of neurophysiology*, 120(2):867–879, 2018.
34. Manvi Goel and Stuart C Mangel. Dopamine-mediated circadian and light/dark-adaptive modulation of chemical and electrical synapses in the outer retina. *Frontiers in Cellular Neuroscience*, 15:647541, 2021.
35. Rolf Herrmann, Stephanie J Heflin, Timothy Hammond, Bowa Lee, Jing Wang, Raul R Gainetdinov, Marc G Caron, Erika D Eggers, Laura J Frishman, Maureen A McCall, et al. Rod vision is controlled by dopamine-dependent sensitization of rod bipolar cells by gaba. *Neuron*, 72(1):101–110, 2011.
36. Jason Jacoby, Amurta Nath, Zachary F Jessen, and Gregory W Schwartz. A self-regulating gap junction network of amacrine cells controls nitric oxide release in the retina. *Neuron*, 100(5):1149–1162, 2018.
37. B Bauer, B Ehinger, and L Åberg. [3 h]-dopamine release from the rabbit retina. *Albrecht von Graefes Archiv für klinische und experimentelle Ophthalmologie*, 215:71–78, 1980.
38. Bernard F Godley and Richard J Wurtman. Release of endogenous dopamine from the superfused rabbit retina in vitro: effect of light stimulation. *Brain research*, 452(1-2):393–395, 1988.
39. MBA Djamgoz and H-J Wagner. Localization and function of dopamine in the adult vertebrate retina. *Neurochemistry international*, 20(2):139–191, 1992.
40. Víctor Pérez-Fernández, Nina Milosavljevic, Annette E Allen, Kirstan A Vessey, Andrew I Jobling, Erica L Fletcher, Paul P Breen, John W Morley, and Morven A Cameron. Rod photoreceptor activation alone defines the release of dopamine in the retina. *Current Biology*, 29(5):763–774, 2019.
41. Charles W Nichols, David Jacobowitz, and Marianne Hottenstein. The influence of light and dark on the catecholamine content of the retina and choroid. *Investigative Ophthalmology & Visual Science*, 6(6):642–646, 1967.
42. Izhak Nir, Rashidul Haque, and P Michael Iuvone. Diurnal metabolism of dopamine in the mouse retina. *Brain research*, 870(1-2):118–125, 2000.
43. Christophe Ribelayga, Yu Cao, and Stuart C Mangel. The circadian clock in the retina controls rod-cone coupling. *Neuron*, 59(5):790–801, 2008.
44. W Wade Kothmann, Stephen C Massey, and John O'Brien. Dopamine-stimulated dephosphorylation of connexin 36 mediates aii amacrine cell uncoupling. *Journal of Neuroscience*, 29(47):14903–14911, 2009.
45. Nan Ge Jin and Christophe P Ribelayga. Direct evidence for daily plasticity of electrical coupling between rod photoreceptors in the mammalian retina. *Journal of Neuroscience*, 36(1):178–184, 2016.
46. Eric M Lasater and John E Dowling. Dopamine decreases conductance of the electrical junctions between cultured retinal horizontal cells. *Proceedings of the National Academy of Sciences*, 82(9):3025–3029, 1985.
47. Yuki Hayashida, Carolina Varela Rodríguez, Genki Ogata, Gloria J Partida, Hanako Oi, Tyler W Stradleigh, Sherwin C Lee, Anselmo Felipe Colado, and Andrew T Ishida. Inhibition of adult rat retinal ganglion cells by d1-type dopamine receptor activation. *Journal of Neuroscience*, 29(47):15001–15016, 2009.
48. Rebekah A Warwick, Alina S Heukamp, Serena Riccitelli, and Michal Rivlin-Etzion. Dopamine differentially affects retinal circuits to shape the retinal code. *The Journal of physiology*, 2023.
49. David S Bredt, Paul M Hwang, and Solomon H Snyder. Localization of nitric oxide synthase indicating a neural role for nitric oxide. *Nature*, 347(6295):768–770, 1990.
50. Ted M Dawson, David S Bredt, Majid Fotuhi, Paul M Hwang, and Solomon H Snyder. Nitric oxide synthase and neuronal nadph diaphorase are identical in brain and peripheral tissues. *Proceedings of the National Academy of Sciences*, 88(17):7797–7801, 1991.
51. R Yamamoto, DS Bredt, SH Snyder, and RA Stone. The localization of nitric oxide synthase in the rat eye and related cranial ganglia. *Neuroscience*, 54(1):189–200, 1993.
52. Ryo Yamamoto, David S Bredt, Ted M Dawson, Solomon H Snyder, and Richard A Stone. Enhanced expression of nitric oxide synthase by rat retina following pterygopalatine parasympathetic denervation. *Brain research*, 631(1):83–88, 1993.
53. Todd A Blute, Bernd Mayer, and William D Eldred. Immunocytochemical and histochemical localization of nitric oxide synthase in the turtle retina. *Visual Neuroscience*, 14(4):717–729, 1997.

54. TODD A BLUTE, MICHAEL R LEE, and WILLIAM D ELDRED. Direct imaging of nmda-stimulated nitric oxide production in the retina. *Visual Neuroscience*, 17(4):557–566, 2000.
55. William D Eldred and Todd A Blute. Imaging of nitric oxide in the retina. *Vision research*, 45(28):3469–3486, 2005.
56. Vikram Palamalai, Ruth M Darrow, Daniel T Organisciak, and Masaru Miyagi. Light-induced changes in protein nitration in photoreceptor rod outer segments. *Mol Vis*, 12: 1543–1551, 2006.
57. Jan J Blom, Todd A Blute, and William D Eldred. Functional localization of the nitric oxide/cgmp pathway in the salamander retina. *Visual Neuroscience*, 26(3):275–286, 2009.
58. Silke Haverkamp and Heinz Wässle. Immunocytochemical analysis of the mouse retina. *Journal of Comparative Neurology*, 424(1):1–23, 2000.
59. Thomas J Giove, Monika M Deshpande, and William D Eldred. Identification of alternate transcripts of neuronal nitric oxide synthase in the mouse retina. *Journal of neuroscience research*, 87(14):3134–3142, 2009.
60. Jan Blom, Tom Giove, Monika Deshpande, and William D Eldred. Characterization of nitric oxide signaling pathways in the mouse retina. *Journal of Comparative Neurology*, 520(18): 4204–4217, 2012.
61. Gerard P Ahern, Vitaly A Klyachko, and Meyer B Jackson. cgmp and s-nitrosylation: two routes for modulation of neuronal excitability by no. *Trends in neurosciences*, 25(10):510–517, 2002.
62. Ari Sitaramayya. Soluble guanylate cyclases in the retina. *Molecular and cellular biochemistry*, 230:177–186, 2002.
63. John Garthwaite. Dynamics of cellular no-cgmp signaling. *Frontiers in Bioscience-Landmark*, 10(2):1868–1880, 2005.
64. Masaru Miyagi, Hirokazu Sakaguchi, Ruth M Darrow, Lin Yan, Karen A West, Kulwant S Aulak, Dennis J Stuehr, Joe G Hollyfield, Daniel T Organisciak, and John W Crabb. Evidence that light modulates protein nitration in rat retina. *Molecular & Cellular Proteomics*, 1(4):293–303, 2002.
65. Douglas T Hess, Akio Matsumoto, Sung-Oog Kim, Harvey E Marshall, and Jonathan S Stamler. Protein s-nitrosylation: purview and parameters. *Nature reviews Molecular cell biology*, 6(2):150–166, 2005.
66. Stephen L Mills and Stephen C Massey. Differential properties of two gap junctional pathways made by aii amacrine cells. *Nature*, 377(6551):734–737, 1995.
67. Scott Nawy and Craig E Jahr. Suppression by glutamate of cgmp-activated conductance in retinal bipolar cells. *Nature*, 346(6281):269–271, 1990.
68. Josefin Snellman and Scott Nawy. cgmp-dependent kinase regulates response sensitivity of the mouse on bipolar cell. *Journal of Neuroscience*, 24(29):6621–6628, 2004.
69. Ryan E Tooker, Mikhail Y Lipin, Valerie Leuranguer, Eva Rozsa, Jayne R Bramley, Jacqueline L Harding, Melissa M Reynolds, and Jozsef Vigh. Nitric oxide mediates activity-dependent plasticity of retinal bipolar cell output via s-nitrosylation. *Journal of Neuroscience*, 33(49):19176–19193, 2013.
70. Alex H Vielma, Adolfo Agurto, Joaquín Valdés, Adrián G Palacios, and Oliver Schmachtenberg. Nitric oxide modulates the temporal properties of the glutamate response in type 4 off bipolar cells. *PLoS One*, 9(12):e114330, 2014.
71. Guo-Yong Wang, Lauren C Liets, and Leo M Chalupa. Nitric oxide differentially modulates on and off responses of retinal ganglion cells. *Journal of neurophysiology*, 90(2):1304–1313, 2003.
72. Guo-Yong Wang, Deborah A Van der List, Joseph P Nemargut, Julie L Coombs, and Leo M Chalupa. The sensitivity of light-evoked responses of retinal ganglion cells is decreased in nitric oxide synthase gene knockout mice. *Journal of Vision*, 7(14):7–7, 2007.
73. Iqbal Ahmad, Trese Leinders-Zufall, Jeffery D Kocsis, Gordon M Shepherd, Frank Zufall, and Colin J Barnstable. Retinal ganglion cells express a cgmp-gated cation conductance activatable by nitric oxide donors. *Neuron*, 12(1):155–165, 1994.
74. Fusao Kawai and Peter Sterling. cgmp modulates spike responses of retinal ganglion cells via a cgmp-gated current. *Visual neuroscience*, 19(3):373–380, 2002.
75. Thomas Euler, Susanne E. Hausselt, David J. Margolis, Tobias Breuninger, Xavier Castell, Peter B. Detwiler, and Winfried Denk. Eyecup scope—optical recordings of light stimulus-evoked fluorescence signals in the retina. *Pflügers Archiv - European Journal of Physiology*, 457:1393–1414, 2009. doi: 10.1007/s00424-008-0603-5.
76. Thomas Euler, Katrin Franke, and Tom Baden. Studying a light sensor with light: Multiphoton imaging in the retina. e. hartveit (ed.). *Multiphoton Microscopy, Neuromethods (Humana, New York, NY)*, 148:225–250, 2019. doi: 10.1007/978-1-4939-9702-2_10.
77. Zhijian Zhao, David A Klindt, André Maia Chagas, Klaudia P Szatko, Luke Rogerson, Dario A Protti, Christian Behrens, Deniz Dalkara, Timm Schubert, Matthias Bethge, et al. The temporal structure of the inner retina at a single glance. *Scientific reports*, 10(1):1–17, 2020.
78. Dominic Gonschorek, Larissa Höfling, Klaudia P Szatko, Katrin Franke, Timm Schubert, Benjamin Dunn, Philipp Berens, David Klindt, and Thomas Euler. Removing inter-experimental variability from functional data in systems neuroscience. *Advances in Neural Information Processing Systems*, 34:3706–3719, 2021.
79. Yongrong Qiu, David A. Klindt, Klaudia P. Szatko, Dominic Gonschorek, Larissa Hoefling, Timm Schubert, Laura Busse, Matthias Bethge, and Thomas Euler. Efficient coding of natural scenes improves neural system identification. *bioRxiv*, page 2022.01.10.475663, 2022. doi: 10.1101/2022.01.10.475663.
80. J Weiss, GA O'sullivan, L Heinze, H-X Chen, H Betz, and H Wässle. Glycinergic input of small-field amacrine cells in the retinas of wildtype and glycine receptor deficient mice. *Molecular and Cellular Neuroscience*, 37(1):40–55, 2008.
81. F Muller, H Wassle, and Th Voigt. Pharmacological modulation of the rod pathway in the cat retina. *Journal of neurophysiology*, 59(6):1657–1672, 1988.
82. Ilya Buldyrev, Theresa Puthussery, and W Rowland Taylor. Synaptic pathways that shape the excitatory drive in an off retinal ganglion cell. *Journal of neurophysiology*, 107(7): 1795–1807, 2012.
83. Amanda J McLaughlin, Kumiko A Percival, Jacqueline Gayet-Primo, and Teresa Puthussery. Glycinergic inhibition targets specific off cone bipolar cells in primate retina. *Eneuro*, 8(1), 2021.
84. J Alexander Bae, Shang Mu, Jinseop S Kim, Nicholas L Turner, Ignacio Tartavull, Nico Kemnitz, Chris S Jordan, Alex D Norton, William M Silversmith, Rachel Prentki, et al. Digital museum of retinal ganglion cells with dense anatomy and physiology. *Cell*, 173(5): 1293–1306, 2018.
85. EJ Chichilnisky. A simple white noise analysis of neuronal light responses. *Network: computation in neural systems*, 12(2):199, 2001.
86. TODD A BLUTE, PAULA VELASCO, and WILLIAM D ELDRED. Functional localization of soluble guanylate cyclase in turtle retina: modulation of cgmp by nitric oxide donors. *Visual neuroscience*, 15(3):485–498, 1998.
87. SEBASTIAN GOTZES, JAN de VENTE, and FRANK MUeLLER. Nitric oxide modulates cgmp levels in neurons of the inner and outer retina in opposite ways. *Visual neuroscience*, 15(5):945–955, 1998.
88. Yongrong Qiu, Zhijian Zhao, David Klindt, Magdalena Kautzky, Klaudia P Szatko, Frank Schaeffel, Katharina Rifai, Katrin Franke, Laura Busse, and Thomas Euler. Natural environment statistics in the upper and lower visual field are reflected in mouse retinal specializations. *Current Biology*, 31(15):3233–3247, 2021.
89. Matías A Goldin, Baptiste Lefebvre, Samuele Virgili, Mathieu Kim Pham Van Cang, Alexander Ecker, Thierry Mora, Ulisse Ferrari, and Olivier Marre. Context-dependent selectivity to natural images in the retina. *Nature Communications*, 13(1):5556, 2022.
90. Larissa Höfling, Klaudia P Szatko, Christian Behrens, Yongrong Qiu, David Alexander Klindt, Zachary Jessen, Gregory S Schwartz, Matthias Bethge, Philipp Berens, Katrin Franke, et al. A chromatic feature detector in the retina signals visual context changes. *bioRxiv*, pages 2022–11, 2022.
91. Thomas Euler, Susanne E Hausselt, David J Margolis, Tobias Breuninger, Xavier Castell, Peter B Detwiler, and Winfried Denk. Eyecup scope—optical recordings of light stimulus-evoked fluorescence signals in the retina. *Pflügers Archiv-European Journal of Physiology*, 457:1393–1414, 2009.
92. Kevin L Briggman and Thomas Euler. Bulk electroporation and population calcium imaging in the adult mammalian retina. *Journal of neurophysiology*, 105(5):2601–2609, 2011.
93. Nicholas M Tran, Karthik Shekhar, Irene E Whitney, Anne Jacobi, Inbal Benhar, Guosong Hong, Wenjun Yan, Xian Adiconis, McKinzie E Arnold, Jung Min Lee, et al. Single-cell profiles of retinal ganglion cells differing in resilience to injury reveal neuroprotective genes. *Neuron*, 104(6):1039–1055, 2019.
94. Katrin Franke, André Maia Chagas, Zhijian Zhao, Maxime JY Zimmermann, Philipp Bartel, Yongrong Qiu, Klaudia P Szatko, Tom Baden, and Thomas Euler. An arbitrary-spectrum spatial visual stimulator for vision research. *elife*, 8:e48779, 2019.
95. Klaudia P Szatko, Maria M Korympidou, Yanli Ran, Philipp Berens, Deniz Dalkara, Timm Schubert, Thomas Euler, and Katrin Franke. Neural circuits in the mouse retina support color vision in the upper visual field. *Nature communications*, 11(1):3481, 2020.
96. William H Press and Saul A Teukolsky. Savitzky-golay smoothing filters. *Computers in Physics*, 4(6):669–672, 1990.
97. Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.
98. Ziwei Huang, Yanli Ran, Jonathan Oesterle, Thomas Euler, and Philipp Berens. Estimating smooth and sparse neural receptive fields with a flexible spline basis. *arXiv preprint arXiv:2108.07537*, 2021.
99. Adrian M Price-Whelan, BM Sipőcz, HM Günther, PL Lim, SM Crawford, S Conseil, DL Shupe, MW Craig, N Dencheva, A Ginsburg, et al. The astropy project: Building an open-science project and status of the v2. 0 core package. *The Astronomical Journal*, 156(3):123, 2018.
100. Hui Zou, Trevor Hastie, and Robert Tibshirani. Sparse principal component analysis. *Journal of computational and graphical statistics*, 15(2):265–286, 2006.

Methods

Animals and tissue preparation. All animal experiments were conducted at the University of Tübingen and were performed according to the laws governing animal experimentation issued by the German Government as well as approved by the institutional animal welfare committee of the University of Tübingen. For all experiments, we used retinae (n=26) from C57Bl/6 J mice (n=14; JAX 000664) of either sex between the age of 4-16 weeks. All animals were kept in the local animal facility and housed under the standard 12h/12h day/night cycle at 22°C and a humidity of 55%.

The following procedures were carried out under very dim red (> 650 nm) light. Before each imaging experiment, the animal was dark-adapted for >1 h, then anesthetized with isoflurane (CP-Pharma) and sacrificed by cervical dislocation. Immediately after, the eyes were enucleated with a dorsal cut as orientation landmark and hemisected in carboxygenated (95% O2, 5% CO2) artificial cerebrospinal fluid (ACSF) solution containing (in mM): 125 NaCl, 2.5 KCl, 2 $CaCl_2$, 1 $MgCl_2$, 1.25 NaH_2PO_4, 26 $NaHCO_3$, 20 glucose, and 0.5 L-glutamine at pH 7.4. Sulforhodamine-101 (SR101, 0.1 µM; Invitrogen) was added to the ACSF to reveal blood vessels and damaged ganglion cell layer (GCL) cells in the red fluorescence channel (75). The carboxygenated ACSF was constantly perfused through the recording chamber at 4 ml/min and kept at ∼36°C throughout the entire experiment. After the dissection, retinae were bulk-electroporated with the synthetic fluorescent calcium indicator Oregon-Green 488 BAPTA-1 (OGB-1; hexapotassium salt; Life Technologies) (92). To electroporate the GCL, the dissected retina was flat-mounted with the GCL facing up onto an AnodiscTM (#13, 0.1 µm pore size, 13 mm diameter, Cytiva), and then placed between two 4 mm horizontal platinum disk electrodes (CUY700P4E/L, Nepagene/Xceltis). The lower electrode was covered with 15 µl of ACSF, while a 10 µl drop of 5 mM OGB-1 dissolved in ACSF was suspended from the upper electrode and lowered onto the retina. Then, nine electrical pulses (∼9.2 V, 100 ms pulse width, at 1 Hz) from a pulse generator/wide-band amplifier combination (TGP110 and WA301, Thurlby handar/Farnell) were applied and then, the electroporated retina on the Anodisc was transferred into the recording chamber, whereby the dorsal edge of the retina pointed away from the experimenter. The retina was left there for 30 min to recover, as well as adapted to the light stimulation by displaying a binary dense noise stimulus (20 x 15 matrix, 40 μm^2 pixels, balanced random sequence) at 5 Hz before the recordings started.

Pharmacology. All used drugs were added to the carboxygenated, perfused ACSF solution 15min prior to the second recording of the GCL scan fields. For the drugs, the respective concentrations were used (in µM): 100 (Z)-1-[N-(2-aminoethyl)-N-(2-ammonioethyl)amino]diazen-1-ium-1,2-diolate (DETA/NO), 10 1H-[1,2,4]oxadiazolo[4,3-a]quinoxalin-1-one (ODQ) and 1 strychnine. The ACSF solution with and without drug application was always kept at ∼36°C.

Two-photon calcium imaging. For the functional calcium imaging experiments, a MOM-type two-photon microscope (designed by W. Denk, MPI, Heidelberg; purchased from Sutter Instruments/Science Products) (75, 76) was employed. The microscope was equipped with a mode-locked Ti:Sapphire laser (MaiTai-HP DeepSee, Newport Spectra-Physics) tuned to 927 nm (ideal wavelength to excite OGB-1), two photomultiplier tubes serving as fluorescence detection channels for OGB-1 (HQ 510/84, AHF/Chroma) and SR101/tdTomato (HQ 630/60, AHF), and a water immersion objective (CF175 LWD×16/0.8W, DIC N2, Nikon, Germany). To acquire images, custom-made software (ScanM by M. Müller and T. Euler) running under IGOR Pro 6.3 for the operating system Microsoft Windows (Wavemetrics) was used and time-lapsed 64 x 64 pixel image scans (∼100 μm^2) at 7.8125 Hz were taken. Routinely, the optic nerve position and the scan field position were recorded to reconstruct their retinal positions. High-resolution images (512 x 512 pixel images) were recorded to support semi-automatic ROI detection.

Light stimulation. For the light stimulation of the retinal tissue, a digital light processing (DLP) projector (lightcrafter (LCr), DPM-E4500UVBGMKII, EKB Technologies Ltd) was used to display the visual stimuli through the objective onto the retina, whereby the stimulus was focused on the photoreceptor layer (94). The LCr was equipped with a light-guide port to couple in external, band-pass filtered green and UV light-emitting diode (LEDs; green: 576 BP 10, F37-576; UV: 387 BP 11, F39-387; both AHF/Chroma). The band-pass filter was used to optimize the spectral separation of mouse M- and S- opsins (390/576 Dualband, F59-003, AHF/Chroma). Both LEDs were synchronized with the scan retracing of the microscope. Stimulator intensity (as photoisomerization rate, 10^3 P^*s^{-1} per cone) was calibrated to range from ∼0.5 (black image) to ∼20 for M- and S-opsins, respectively. A steady illumination of ∼10^4 P^*s^{-1} per cone was present during the scan recordings due to the two-photon excitation of photopigments (75, 76).

In total, 3 types of light stimuli were used for the imaging of calcium in the GCL: (1) full-field chirp stimulus (700 µm ⌀; see details here ref. (5)), (2) bright moving bar (0.3 x 1 mm) at 1 mm s^{-1} in eight directions to probe direction and orientation selectivity, and (3) random binary noise with a checkerboard grid of 20 x 15 checks and a check size of 40 µm at 5 Hz for 5 min to map receptive fields. Before each stimulus was presented, the baseline was recorded after the laser started scanning for at least 30 s to avoid immediate laser-induced effects on the retinal activity (75, 76, 95).

Data Analysis. Image extraction and semi-automatic region-of-interest (ROI) detection were performed using Igor PRO 8. All analyses were organized and performed in a custom-written schema using DataJoint for Python http://datajoint.github.io/; D. Yatsenko, Tolias lab, Baylor College of Medicine.

Preprocessing. After the Ca^{2+} were extracted from individual ROIs, as described elsewhere (5, 95), the raw traces were detrended by subtracting a smoothed version of the trace from the raw one. Detrending was necessary to remove slow drifts in the signal that were unrelated to the light-induced response. The smoothed version was the trace was computed by applying a Savitzky-Golay filter (96) of 3^{rd} polynomial order and a window length of 60 s using the Python SciPy implementation $scipy.signal.savgol\_filter$ (97).

$$r_{detrend} = r_{raw} - r_{smooth}$$

Lastly, the baseline activity (mean of the first eight samples) was subtracted, the mean activity $r(t)$ was computed and then traces were normalized such that:

$$\max_t(|r(t)|) = 1$$

Inclusion criterion. To include reliable cell responses for the performed analyses, two consecutive quality filtering steps were applied. At first, the response quality criterion, also termed quality index (QI), was computed for the moving bar ($QI_{MB} > 0.6$) and full-field chirp ($QI_{chirp} > 0.45$). Cells that passed either one of these two QIs in both recording conditions were included, otherwise they were discarded in the following analyses. As in Baden et al. (5), the quality index is defined as follows:

$$QI = \frac{Var[\langle C \rangle_r]_t}{\langle Var[C]_t \rangle_r}$$

where C is the T by R response matrix (time samples by stimulus repetitions) and $\langle\rangle_x$ and $Var[]_x$ denote the mean and variance across the indicated dimension x, respectively. As a second step, cells were assigned to a RGC group using the RGC classifier, which returned the RGC group index and a confidence score (i.e., assignment probability to the predicted RGC group by the random forest classifier (79)). Only cells that were assigned to one of the RGC groups (i.e., RGC index 1 - 32) were included, whereby cells assigned to a displaced AC group (i.e., RGC index 33 - 46) were rejected. Cells that exceeded the confidence score threshold of > 0.25 were included.

Suppression index. For each cell, the suppression index (SI) was measured by comparing the (absolute) negative area under the curve (AUC_{neg}) of the chirp and moving bar responses with the total area under the curve (AUC_{neg}+AUC_{pos}) of the entire response trace. For the absolute AUC_{neg}, the response was clipped for value <0.

$$SI = \frac{|AUC_{neg}|}{|AUC_{neg}| + |AUC_{pos}|}$$

On-Off Index. The On-Off index was computed as

$$OOI = \frac{\langle r_{on} \rangle_t - \langle r_{off} \rangle_t}{\langle r_{on} \rangle_t + \langle r_{off} \rangle_t}$$

where r_{on} and r_{off} are defined as the separated time components of the moving bar response into its On- and Off-component. For each component, we computed the mean value of the discrete differences along the time axis clipped between 0 and 1 to estimate if there is a response to the particular feature.

Classification of functional retinal ganglion cell types. For the functional classification of RGC types, we used a previously published RGC classifier (79). The classifier, which uses a random forest classifier, was trained, validated and tested on previously published RGC type responses (5). As input to the classifier, we used the responses to the standard set of stimuli, i.e., full-field chirp and moving bar, as well as soma sizes (separates alpha and non-alpha types) and the p-values of the permutation test for direction selectivity (separates DS and Non-DS types). For every cell, the RGC classifier outputs its type index and the confidence scores for all 46 types. The confidence score, as described in 'Inclusion criterion' was used as quality criterion.

Receptive field estimation. We mapped receptive fields (RFs) of RGCs using the RF python toolbox *RFEst* (98), following the procedure in (5) with few modifications. The binary dense noise stimulus (20 × 15 matrix, 40 µm^2 pixels, balanced random sequence; 5 Hz) was centered on the recording field. We computed the temporal gradients of the Ca^{2+} signals from the detrended traces and clipped negative values:

$$\dot{c} = \max(0, \dot{r}_{detrend})$$

The stimulus $S(t)$ and the clipped temporal gradients c were upsampled to 10 times the stimulus frequency to compute the gradient-triggered average stimulus:

$$\boldsymbol{F}(x,y,\tau) = \int_{t=0}^{T} \dot{c}(t)\boldsymbol{S}(x,y,t+\tau)$$

where $\boldsymbol{S}(x,y,t)$ is the stimulus, τ is the lag ranging from approximately -0.20 to 1.38 seconds, and T is the duration of the stimulus. We smoothed these raw RF using a 5 × 5 pixel and 1 pixel standard deviation Gaussian window for each lag. Then we decomposed the RF into a temporal ($F_t(\tau)$) and spatial ($\boldsymbol{F}_s(x,y)$) component using singular value decomposition and scaled them such that $\max(|F_t|) = 1$ and $\max(|\boldsymbol{F}_s|) = \max(|\boldsymbol{F}|)$. For each spatial RF $\boldsymbol{F}_s$, we fit a 2D Difference of Gaussians using the python package *astropy* (99). The mean and covariance matrices of the center and surround Gaussian fits were tied, except for a linear scaling of the covariance matrix. We defined the polarity $p \in \{-1, 1\}$ of the spatial RF as the sign of the model fit at its mean. Next, we computed the center RF of the spatial RF as $\boldsymbol{F}_s^c = \max(0, p \cdot \boldsymbol{F}_s)$. The surround index was computed as:

$$RF_{surround} = \frac{\sum_{x,y}(\boldsymbol{F}_s - \boldsymbol{F}_s^c)}{\sum_{x,y}(|\boldsymbol{F}_s|)}.$$

To measure the center RF size, we fit a 2D Gaussian to the center RF, with the mean fixed to the one obtained from the Difference of Gaussians fit. The area covered by two standard deviations of this Gaussian fit was used as the RF size.

Functional clustering. The functional clustering was based on a similar approach as in (5). The clustering was only applied on RGC types previously classified as G_{32} and only recorded in Ctrl 1. First, visual features from the full-field chirp and moving bar Ca^{2+} responses were extracted using sparse principal component analysis (100). After optimizing the alpha parameter for each stimulus, each cell's dimensionality was reduced to 30 features, whereby the chirp covered 21 and the moving bar 9 features. Alpha was optimized in a way that every part of each stimulus was represented by one feature to increase interpretability. Each feature was standardized across cells before clustering. Then, the features were used to cluster the cells using a Mixture of Gaussian model. The ideal number of clusters was chosen based on the cross-validated Bayesian Information Criterion. Additionally, cluster coherence was computed and validated using intra- and inter-cluster correlation, as well as the influence of potential batch effects, i.e., a single cluster originates from a single retina or scan field, but is found across several ones. Then, the model was used to predict cell types labels. Finally, cluster 3 showed a high signal-to-noise ratio, thus cells were re-clustered, which originated in three clusters, whereby two showed high variability in their chirp and moving bar responses. These cells were discarded in the further analysis to clean this cluster from potential contamination.

Statistical analysis. To quantify the differences between traces, a Shaprio-Wilk test was used to test for normality and then either the two-sided t-test (if normally distributed) or the non-parametric Wilcoxon signed-rank test (Mann-Whitney U Test). To determine α, Bonferroni correction was used, depending on the number of tests performed. To test the difference between traces against zero, we either used the t-test or the non-parametric Wilcoxon signed-rank test, depending on the distribution. A one-sampled T-test was performed to test the mean against a population mean of zero in order to quantify if the mean difference diverges from zero. For the statistical comparison of the suppression index between conditions per cluster, the non-parametric Kruskal-Wallis one-way analysis of variance and post-hoc Dunnett's test and Bonferroni correction to determine the statistical significance between conditions was used.

Data and code availability. Data, light stimuli, as well as all custom analyses including code and notebooks to reproduce analyses will be made available at https://github.com/eulerlab upon journal publication.

Author contributions. D.G., T.S.K., Z.Z., T.E. and T.S. designed the study; D.G. performed imaging experiments, supported by R.A. and T.S.K.; D.G. performed pre-processing; D.G. analysed the data with the help of J.O.; Z.Z. helped to maintain the experimental setup; D.G. wrote the manuscript with the help from T.E., T.S., J.O.

Supplementary Material

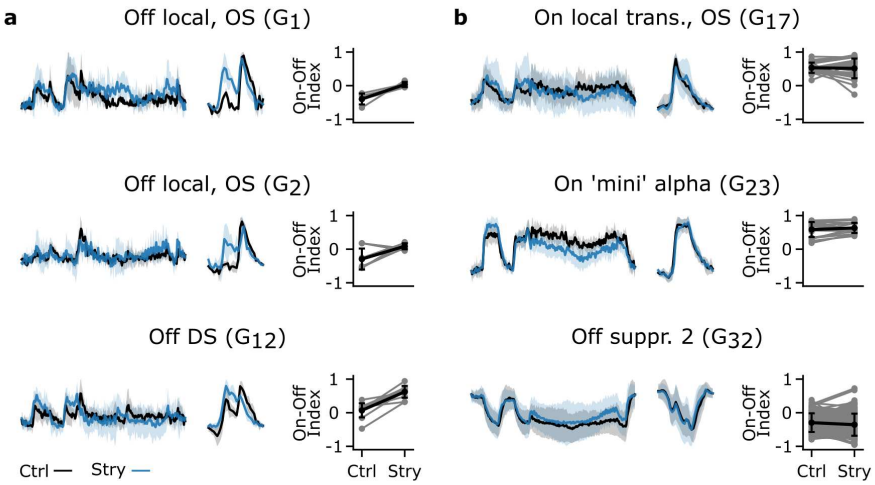

Fig. S1. Testing the effects of strychnine on different RGC type responses. **(a)** Left: Representative mean calcium responses of paired RGC types between Ctrl (black) and Strychnine-application (blue) (top: G_1; middle: G_2; bottom: G_{12}) exhibiting Off-responses. Right: Corresponding On-Off indices of paired RGCs for the Ctrl and Strychnine condition. **(b)** Left: As **(a)**, representative mean calcium responses of paired RGC types between Ctrl (black) and Strychnine-application (blue) (top: G_{17}; middle: G_{23}; bottom: G_{32}) exhibiting mainly On-responses and a suppressed-by-contrast component. Right: Corresponding On-Off indices of paired RGCs for the Ctrl and Strychnine condition.