Summer 8-15-2017

Robust Odorant Recognition in Biological and Artificial Olfaction

Nalin Katta

Washington University in St. Louis

Recommended Citation

Katta, Nalin, "Robust Odorant Recognition in Biological and Artificial Olfaction" (2017). *McKelvey School of Engineering Theses & Dissertations*. 308.
https://openscholarship.wustl.edu/eng_etds/308

WASHINGTON UNIVERSITY IN ST. LOUIS

School of Engineering & Applied Science
Department of Biomedical Engineering

Dissertation Examination Committee:
Baranidharan Raman, Chair
Parag Banerjee
Jiamin Cui
Timothy Holy
Daniel Moran
Srikanth Singamaneni

Robust Odorant Recognition in Biological and Artificial Olfaction

by
Nalin Katta

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2017
St. Louis, Missouri

Table of Contents

List of Figures

Acknowledgments

Science is not done in a vacuum, but rather is a group effort. I would like to first thank all my collaborators, without whom the work in this thesis would not be possible. From Aim 1: Dr. Kurt Benkstein, Dr. Douglas C. Meier, and Dr. Steven Semancik. From Aim 2: Matthew O'Neill, Lijun Zhang, and Rishabh Chandak. From Aim 3: Chad Schaber, Lucy Bollinger, Dr. Indi Trehan, and Dr. Audrey R. Odom John. Working with you has been an honor and pleasure.

I would also like to thank Debajit Saha, with whom I have worked with for my entire time in the Raman Lab. Your mentoring and friendship have been invaluable.

I thank my committee, Dr. Srikanth Singamaneni, Dr. Parag Banerjee, Dr. Tim Holy, Dr. Dan Moran, and Dr. Jiamin Cui. Your thoughtful questions and feedback have improved this work tremendously.

Finally, I would like to thank my research advisor, Dr. Baranidharan Raman. Your guidance and mentorship have made me a better scientist.

Nalin Katta

Washington University in St. Louis

August 2017

Dedicated to my parents.
Everything I have accomplished is a result of your love and support.

ABSTRACT OF THE DISSERTATION

Robust odorant recognition in biological and artificial olfaction

by

Nalin Katta

Doctor of Philosophy in Biomedical Engineering

Washington University in St. Louis, 2017

Professor Baranidharan Raman, Chair

Accurate detection and identification of gases pose a number of challenges for chemical sensory systems. The stimulus space is enormous; volatile compounds vary in size, charge, functional groups, and isomerization among others. Furthermore, variability arises from intrinsic (poisoning of the sensors or degradation due to aging) and extrinsic (environmental: humidity, temperature, flow patterns) sources. Nonetheless, biological olfactory systems have been refined over time to overcome these challenges. The main objective of this work is to understand how the biological olfactory system deals with these challenges, and translate them to artificial olfaction to achieve comparable capabilities. In particular, this thesis focuses on the design and computing mechanisms that allow a relatively simple invertebrate olfactory system to robustly recognize odorants even though the sensory neurons inputs may vary due to the identified intrinsic, or extrinsic factors.

In biological olfaction, signal processing in the central circuits is largely shielded from the variations in the periphery arising from the constant replacement of older olfactory sensory neurons with newer ones. Inspired by this design principle, we developed an analytical method where the operation of a temperature programmed chemiresistor is treated akin to a mathematical

input/output (I/O) transform. Results show that the I/O transform is unique for each analyte-transducer combination, robust with respect to sensor aging, and is highly reproducible across sensors of equal manufacture. This enables decoupling of the signal processing algorithms from the chemical transducer, and thereby allows seamless replacement of sensor array, while the signal processing approach was kept a constant. This is a key advance necessary for achieving long-term, non-invasive chemical sensing.

Next, we explored how the biological system maintains invariance while environmental conditions, particularly with respect to changes in humidity levels. At the sensory level, odor-evoked responses to odorants did not vary with changes in humidity levels, however, the spontaneous activity varied significantly. Nevertheless, in the central antennal lobe circuits, ensembles of projection neurons robustly encoded information about odorant identity and intensity irrespective of the humidity levels. Interestingly, variations in humidity levels led to variable compression of intensity information which was carried forward to behavior. Taken together, these results indicate how the influence of humidity is diminished by central neural circuits in the biological olfactory system.

Finally, we explored a potential biomedical application where a robust chemical sensing approach will be immensely useful: non-invasive assay for malaria diagnosis based on exhaled breath analysis. We developed a method to screen gas chromatography/mass spectroscopy (GC/MS) traces of human breath and identified 6 compounds that have abundance changes in malaria infected patients and can potentially serve as biomarkers in exhaled breath for their diagnosis. We will conclude with a discussion of on-going efforts to develop a non-invasive solution for diagnosing malaria based on breath volatiles.

In sum, this work seeks to understand the basis for robust odor recognition in biological olfaction and proposes bioinspired and statistical solutions for achieving the same abilities in artificial chemical sensing systems.

Chapter 1: Introduction

Chemical sensory systems face several challenges in accurately detecting and identifying gases. The stimulus space is enormous; volatile compounds vary in size, charge, functional groups, and isomerization among others. Efforts to quantify the size of this space have calculated 166.4 billion possible compounds of 17 or fewer C, N, O, S, and halogen atoms [1]. The number would vastly increase as the possible number or type of atoms increases. Further variability arises from environmental factors such as air flow, temperature, and humidity. The biological chemical sensing systems, however, have been refined over time to overcome these challenges. Strikingly, features of biological olfaction are largely conserved across phyla [2], suggesting that biology has come up with one solution for these challenges. The work in this thesis will investigate a simple olfactory system of the american locust *Schistocerca Americana*.

1.1 The locust olfactory system

The biological olfactory system is a finely tuned machine that has been refined by evolution over eons. It serves many purposes including, but not limited to, foraging for food, threat or predator detection, finding mates, as well as for receiving signals and communication. While there are subtle differences in the system and design across species, olfaction is widely conserved across phyla [2]. The work in this thesis uses the American locust, *Schistocerca americana*, as a model system. Locusts have been a popular system for olfaction research for decades. Their use has a number of key advantages: they are well characterized, they are of a manageable size, and their sensory organs as well as central nervous system are readily accessible. Furthermore, recent studies have shown the ability and utility of this system in behavioral assays [3, 4]. In this section, we will examine the locust olfactory system and take a look at its physiology and function. We will begin with the sensory organ for locusts, the antenna. Then proceed to the primary processing center, the antennal lobe. And finally we will examine the mushroom bodies and other downstream centers involved in olfactory processing.

1.1.1 The antenna

The locust antenna is composed of several flagellar segments called annuli. On these segments are hundreds of small cone-like structures called sensilla, which house the olfactory receptor neurons (ORNs). ORNs are sensory neurons through which airborne odorants are first transduced into electrical signals and transmitted to the brain. Each odor activates a combinatorial response across the array of ORNs, and each ORN in turn responds to several ligands. Both the number of spikes (firing rate) and the temporal pattern of spike trains tend to be odor specific [5]. Typical ORNs send a dendrite into the sensilar projection, while the cell body is located at the base. The ends of these dendrites have small cilia with pores for receiving odorant molecules [6]. ORNs express a specific receptor from a large multifamily receptor gene [7, 8]. While each ORN only expresses one receptor, each receptor is capable of detecting a variety of odorants and concentrations [9].

There are four main types of sensilla present on locust antenna (Figure 1.1), basiconica, trichodea, coeloconica, and chaetica. They each have a different function, different distributions across different development stages and antennal segments, as well as house different numbers of ORNs. In general, they house ORNs whose axons are bundled in an antennal nerve and synapse directly to the antennal lobe. The specific details for many of these features can be found in the fine work by Ochieng' et al. [10], here I will present an overview of the structure of the antenna and its components.

Sensilla basiconica are the primary sensilla type involved in odorant detection. These are the largest sensilla type. They are 5 μm in diameter at the base and have an average length of 16 μm. They house the most number of ORNs, typically between 20-50 ORNs per sensilla [10, 11]. They have been shown to respond to general odors as well as aggregation pheromones [12]. Basiconica are also the most common sensilla type in both adult and immature locusts (nymphs) [10]. These correspond to type-A sensilla basiconica in *Locusta migratoria* reported in works by Chapman and Greenwood [13, 14].

Sensilla trichodea are smaller, thin sensilla. They have a base diameter of 3 μm and a length of 8-10 μm and only contain 1-3 ORNs [10]. Some works have reported that there is only one apical pore on these sensilla, suggesting that they must be contact chemoreceptors [13, 14]. Others have reported that there are several pores [10], suggesting that there might be an alternate role for these sensilla as general olfactory receptors or pheromone receptors [12, 15]. These correspond to type-B sensilla basiconica in *Locusta migratoria* reported in works by Chapman and Greenwood [13, 14].

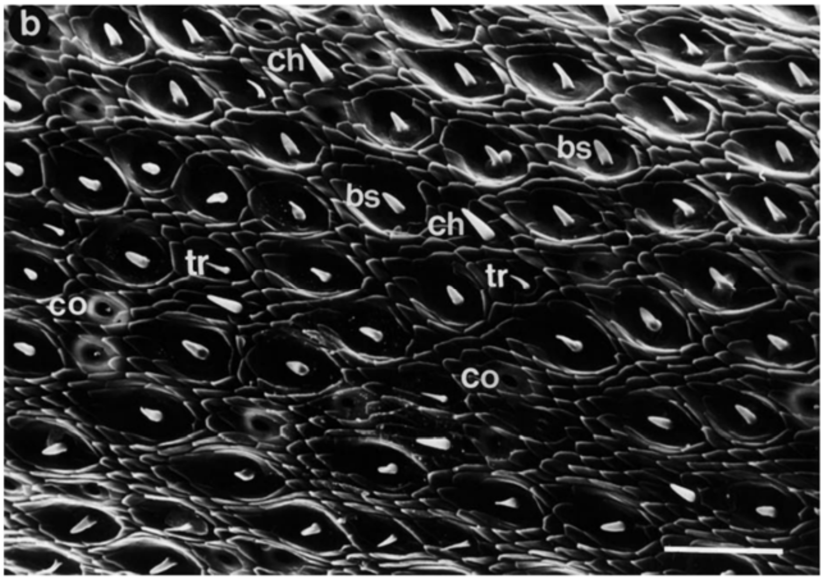


Figure 1.1. SEM image of antenna from adult locust, *S. gregaria*. The four different sensilla types are marked on this segment; bs = basiconica, tr = trichodea, co = coeloconica, ch = chaetica. Scale bar indicates 50 µm. Image reproduced as is from [10].

Sensilla coeloconica are the second most common sensilla on the locust antenna. They do not extrude externally like the other sensilla types, but rather are located in small circular pits 5 µm in diameter. Coeloconica are found in two main varieties and are among the more interesting sensilla types. One variety is double-walled, has radial pores throughout its surface [16], and contains three ORNs. The second has no pores and contains four ORNs [10]. This second type of sensilla is responsible for detecting changes in humidity and temperature and does not respond to the presence of odorants [17]. The first type, on the other hand, has been shown to respond strongly to caproic acid, butyric acid, and hexanal [6, 18, 19]. Interestingly, these two types are un-evenly distributed across the antenna. The first type represents 96% of the coeloconica sensilla present, while the second represents only about 4% [17].

Sensilla chaetica are between basiconica and trichodea in size and shape. They have a diameter of 4µm at the base and are 13-16 µm long. While the exact function of these sensilla remains unclear, their physical properties suggest possible functional roles. They are based in a flexible socket, suggesting they might be used as mechanoreceptors [10]. Furthermore, they only have one apical pore and no pores along the external wall, contributing to the notion that they are not meant to detect odorants in the air.

1.1.2 The antennal lobe

The locust antennal lobe is the insect equivalent of the mammalian olfactory bulb. Here, inputs from the sensory neurons, ORNs, are integrated and processed before being passed downstream to higher level centers (Figure 1.2). The antennal lobe primarily consists of two neuron types.

The first neuron type we will discuss are projection neurons (PNs). PNs receive inputs from 50 thousand ORNs in the antenna and ‘project’ them to two areas of the brain, the mushroom body and lateral horn. There are 830 PNs in a locust antennal lobe, 10-15% of which are activated by any one olfactory stimulus [20-23]. PN cell bodies are all located on the surface of the antennal lobe, while the bulk of the volume is made up of thousands of small bundles of neuropils called glomeruli [24]. Dendrites from PNs extend into glomeruli and receive inputs from the antenna. Several different PNs branch into the same glomerulus, and one PN branches to 10-20 glomeruli [20, 24]. Inputs from ORNs are also collected in the glomeruli. Interestingly, as shown in both vertebrate and invertebrate systems, ORNs expressing the same receptor synapse into the same glomeruli [25, 26].

The second type of neurons, local neurons (LNs), are an integral part of the antennal lobe computations and stimulus processing. LNs are GABAergic inhibitory neurons which synapse with PNs and other LNs in the antennal lobe [20]. There are approximately 300 LNs, all of which extensively arborize throughout the antennal lobe[22, 27]. The antennal lobe network is driven to oscillatory synchrony during an odorant stimulus to help drive signaling and odor identification in downstream centers. Blocking LN activity disrupts that synchronization [27]. LNs are unique neurons in two ways. First, they do not have axons; instead they arborize the antennal lobe entirely with dendritic projections. Second, they do not fire full action potentials, but rather have spikelets, which are smaller more wavelike changes in electrical potential.

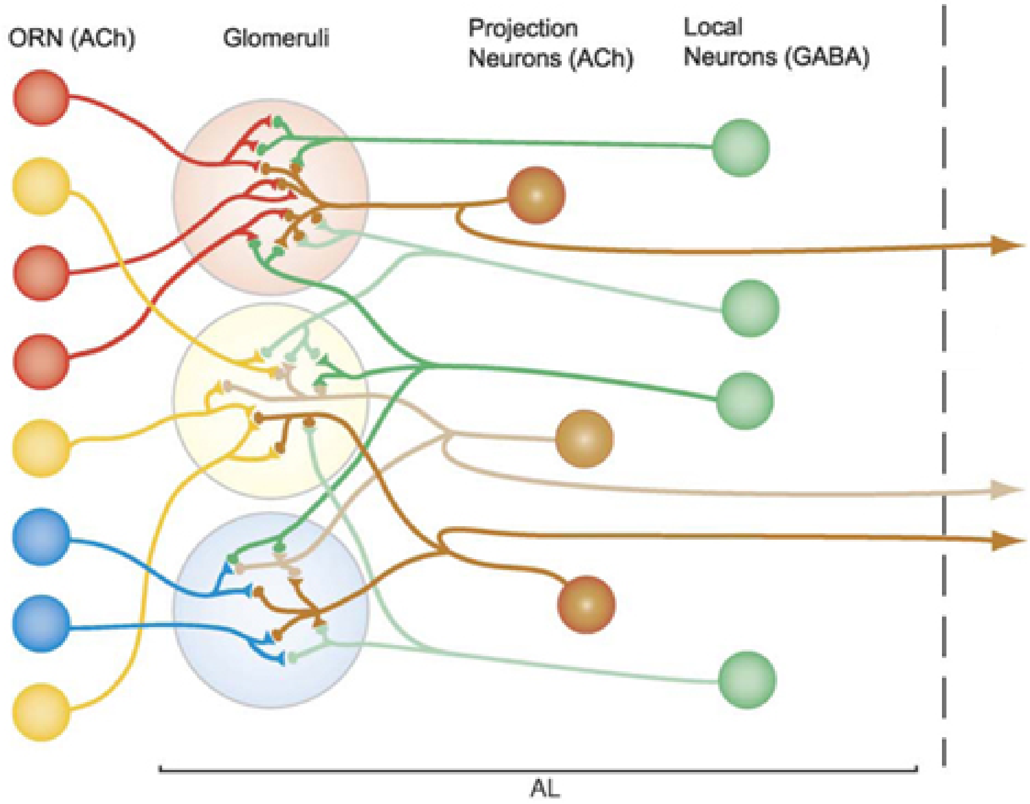


Figure 1.2. Schematic of locust olfactory circuit from antenna to antennal lobe. ORNs expressing the same receptor are colored the same and synapse to the same glomeruli. PNs (brown) and LNs (green) can synapse in several glomeruli. Figure adapted from [9].

1.1.3 The mushroom body

The mushroom body is a region of the insect brain where inputs from PNs in the antennal lobe are received and undergo further processing. These areas are generally accepted to be responsible for memory and learning [28]. Axons from PNs branch widely in the mushroom bodies and synapse with the primary cells of this region, Kenyon cells (KCs).

There are nearly 50,000 KCs in each mushroom body. Each one receives inputs from 10-20 PNs and each PN connects to nearly 600 KCs [21, 28]. Given the number of PNs that synapse with each KC, one would expect them to respond to a wide variety of odorants. The opposite, however, is true: an individual KC will respond to very few odorants in a specific concentration range [29, 30]. Furthermore, stimulus-evoked action potentials in the KCs are sparse, firing only 1-2 spikes in response to an odor stimulus and they have almost no baseline activity [29, 30].

Several factors account for the sparse activity in KCs. One is that KCs have high thresholds for firing caused by the inhibitory giant GABAergic neuron (GGN). This neuron is bi-directionally connected to all KCs and is broadly responsive to any odorant stimulus [30, 31]. It thus depresses the activity of KCs. A second factor is that, in addition to excitatory inputs from PNs, KCs also receive phase-delayed inhibition from neurons in the lateral horn, another region of the insect brain [9, 20, 27, 28]. Therefore, in order to initiate action potentials in KCs, a group of PNs must execute synchronized firing during the short period when the KCs are not being inhibited by the lateral horn neurons.

1.2 Artificial chemical sensing

Methods to detect and identify gases first began in the early 1940's with chromatographic techniques and followed advances in resonant devices, semi-conductors, conducting polymers, and fiber optics (Figure 1.3). This section, briefly reviews the history of gas sensor development and the primary devices used. Two sensor technologies used in this dissertation are gas chromatography/mass spectroscopy and solid-state sensors.

1.3 Gas chromatography

The first mention of chromatography being applied to gases was suggested by Martin and Synge in their 1941 paper advancing chromatography theory [32]. This Nobel Prize winning work was later developed by James and Martin into a working method for a gas chromatogram [33]. Gas chromatography (GC) primarily relies on the general chromatography principle that the separation of any mixture requires a mobile phase which moves the mixture down or through a stationary phase (i.e., a column) which has different partition coefficients or affinities to the components of the mixture [33-35]. Essentially, different components of a mixture travel through a column at different speeds, depending on their affinity for the column materials. Originally, GC started with large columns, meters in length and millimeters in inner diameter (I.D.), 'packed' with particles coated in a stationary phase material [33, 35]. But the idea of using am internally coated capillary column, first implemented by Golay in 1958 [35], has now become the standard for modern gas chromatography.

DEVICE CATEGORY	**SENSING PRINCIPLE**
COLUMN BASED	Gas chromatography[
SPECTROSCOPIC[9]	Mass spectroscopy Ion mobility spectroscopy
ELECTRICALLY-BASED/CHEMIRESISTANCE[10]	Metal oxide films Field effect transistors Conducting polymers
GRAVIMETRIC[11]	Surface acoustic wave Bulk acoustic wave Thin-film bulk acoustic wave Quartz crystal microbalance
OPTICAL[12, 13]	Colorometric arrays Fluorescence Reflectometric interference spectroscopy Infrared spectroscopy
BIOLOGICALLY-BASED	Olfactory receptors expressed in cells Olfactory receptors bound to Field effect transistors[14]

Figure 1.3. Current gas sensor device categories and sensor modalities.

1.3.1 Modern gas chromatography

Modern GC still relies on a column as the primary mechanism to elute a gas. Unlike the first gas chromatograms, which used tubes packed with the stationary phase, modern GCs instead use capillaries which are coated with the stationary phase along their inner diameter. While packed columns had dimensions of 1-5 mm in inner diameter (I.D.) and up to 5 m in length, capillary columns are much thinner and longer, less than 1 mm in I.D. and up to 100 m in length [35]. Furthermore, as with all methods of chromatography, GC separates the components of gases, but lacks a specific way to identify the components. For identification, a detector, must be attached in line with the column output. There are a number of detectors available for GC, operating on principles as diverse as flame ionization, thermal conductivity, and atomic emission detection. The detector can be chosen based on selectivity, limits of detection, and range of detection [36].

1.3.2 Applications areas for gas chromatography

Since its invention in 1952, GC has become the gold standard for gas identification. While its initial use was primarily in the oil industry for identifying components of different petroleum products, the uses have greatly expanded since. Benefiting from the variety of techniques and advancements for GC now, can be used for such purposes as detecting airborne volatiles and chemical composition analysis of solids (e.g., soil samples, plant matter, pharmaceuticals). Furthermore, interest has expanded in miniaturization and portability [35], opening the door for even further applications.

1.4 Solid-state sensors

There are several other classes of chemical vapor sensing devices. Commonly used devices include quartz crystal microbalances, surface acoustic wave devices, and chemiresistive films (conducting polymers, nanoparticle monolayers, and metal-oxides) [37]. All of these have different signaling modalities, but share a common principle: A volatile organic compound (VOC) interacts with the surface of the sensor, resulting in a change in a measureable property.

In the case of chemiresistive sensors, the conductance/resistance is changed. The work in this thesis will focus on chemiresistive film-based gas sensing devices, specifically metal-oxide (MOX) films.

1.4.1 Metal-oxide films

The usage of MOXs in gas sensing applications was initiated by Seiyama et al. in 1962 when they reported the use of a ZnO based gas sensor [38, 39]. Since then, several other MOXs have come in use for gas sensing applications, including: ZnO, TiO_2, and, one of the most popular, SnO_2 (Figure 1.4). As mentioned above, the conductance of these films change as the result of reactions in which the concentration of surface oxygen changes [40, 41] and bulk reactions, that rely on point defects in the MOX crystal [40].

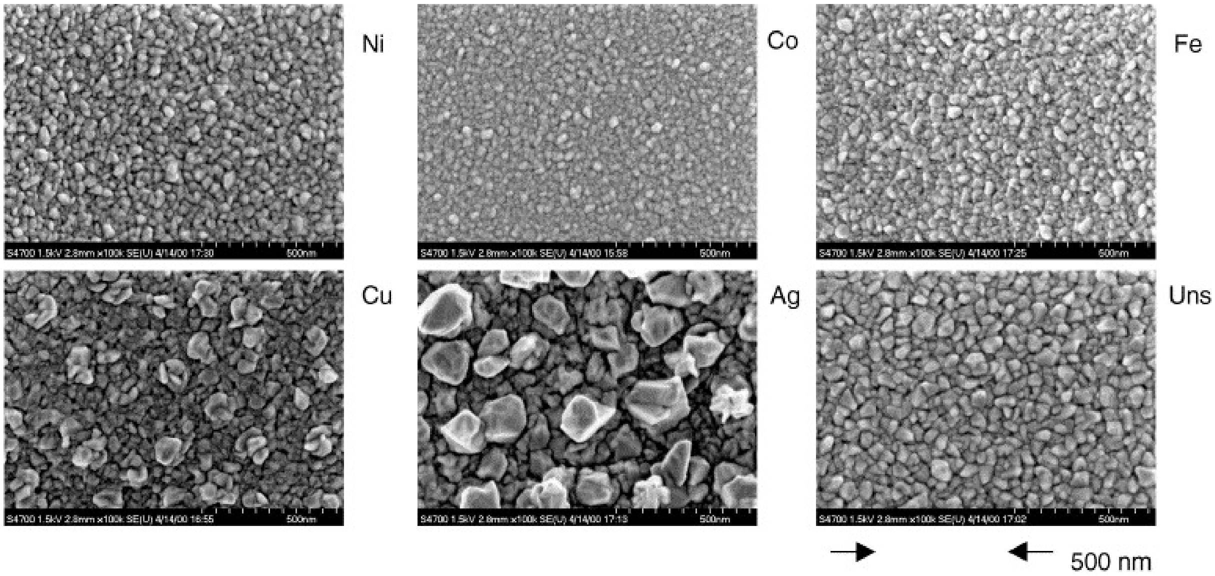


Figure 1.4. SEM micrographs of SnO_2 films. Films were deposited using different metal seed layers in a chemical vapor deposition technique. Note that seed metal Uns (atomic number 107) is now know as Bohrium. Figure as is from [42].

Downsides of MOX-based sensors include low selectivity and required heating to high temperatures (> 200°C) [43, 44], mostly to overcome the interference of environmental water vapor [45]. The need for heating results in increased device complexity and energy requirements. There is, however, a silver lining. Changes in film temperature can change reaction kinetics and film conductance, among other material properties [42]. The resulting optimal temperature zones for different VOCs, impart greater selectivity to the sensor.

To reduce energy expenditures for heating, researchers have sought to miniaturize the heating and sensing elements. Modulating temperature has been one of the techniques used to increase the dimensionality of response from a sensor. Through CMOS (see Figure 2.4) and other miniaturization techniques, devices with rapid thermal time constants and individually addressable sensors can rapidly cycle through temperatures [42]. This rapid cycling allows for the use of “pseudo-sensors”, sensors which are based on the same material, but have different reaction properties at different operating temperature [46, 47].

1.5 Limitations of artificial olfaction

In this work, I will focus on addressing three current challenges in artificial olfaction, or electronic noses, which are solved in biological olfaction. The first aim focuses on the problem of drift, or deviation of the response over time, and finding a means to overcome it in almost any chemical sensor. The second aim focuses on overcoming environmental variability, specifically variations in humidity, which also cause changes in sensor responses. In the third and final aim, will explore sensitivity and selectivity of detection.

1.5.1 Sensor drift

Drift is an issue faced by almost all chemical sensors that severely limits their long-term use [46, 48-54]. It is thought to be primarily an effect of either aging or poisoning of the sensing film [55], which profoundly alters absolute transducer measurements, making these measures unreliable for long term analyte identification. Aging is caused by intrinsic changes to the sensor over time and with use [40]. For chemiresistive films, it can be brought about by changes in the internal bonds and material properties of the film. Poisoning, on the other hand, is the result of extrinsic changes to the chemical film caused by permanent or semi-permanent binding of test gases or by environmental conditions. Because of these effects, there is a long-standing, unmet need for portable, accurate, and precise chemical sensors which remain viable for extended periods of operation.

Natural systems, however, have developed mechanisms to overcome the effects of aging and poisoning. One such mechanism is the turnover of ORNs, the individual sensors that are responsible for primary chemical detection. It is well known that olfactory sensory neurons which transduce chemical cues into electrical signals are constantly regenerated[56] and integrated into the circuit in an unsupervised fashion[57-59]. Regeneration of olfactory receptors has been shown across phyla in both simple organisms, such as insects[23], and higher organisms, such as mammals [56, 57, 60].

A second mechanism is the remarkable ability of regenerated ORNs to re-innervate the brain during regular regeneration of the olfactory sensors and primary signal processing cells during regeneration after injury [23, 57, 61] or apoptosis. Hence, the central circuitry in the brain can remain largely unaffected by changes to the peripheral sensory systems, allowing for long term viability of the system.

1.5.2 Sensor environmental invariance

Another difficulty for chemical sensors is a constantly fluctuating environment. In practical applications, sensors will be deployed in areas where conditions regularly change; the background gas will have varying components and concentrations, air flows will change in speed and direction, temperatures will fluctuate, and humidity levels will change. Humidity, in particular, has drastic effects on a chemical sensor signal [45]. All of these things are affected by generally uncontrollable parameters such as time of day, season, etc. Nevertheless, to be successful, detection and recognition of targets must remain invariant to these fluctuations.

Biological olfaction has excellent invariance to environmental variation. Your ability to smell does not change much with temperature fluctuations between summer and winter, nor does it change much between rainy and dry days. What enables this invariance in natural systems?

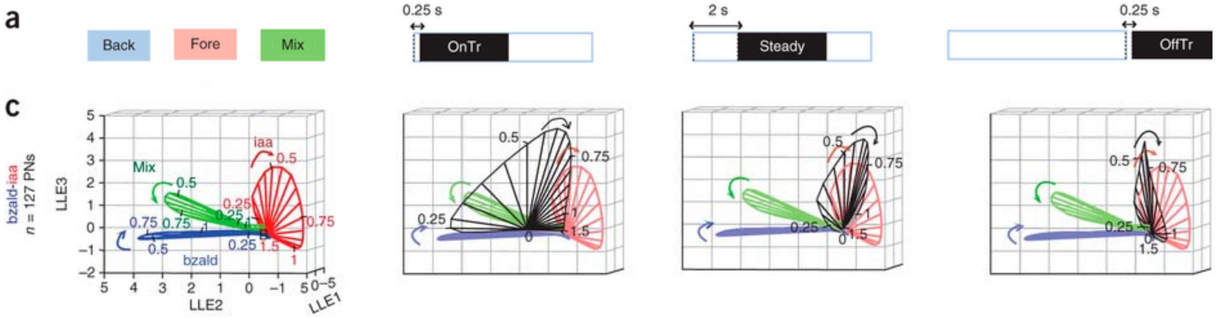


Figure 1.5. Spatiotemporal code encoding of stimulus sequences. Neural trajectory shifts during the stimulus sequence. Depending on the offset between background and target odorant, or foreground, the extent of pattern match changes, yet match with the target odorant remains. Figure adapted from [62].

Several studies have been performed assessing the olfactory system's ability to pick out a target in changing or complicated background composition [4, 63-65]. Studies that have examined the mechanisms for allowing this invariance have found that it comes primarily from the processing performed by the olfactory system [4, 63]. In locusts exposed to mono-molecular odorants as well as their binary mixtures, Saha et al. found that the spatiotemporal code encoding odorants shifted depending on the background (Figure 1.5). But, their results showed that a partial match in the spatiotemporal code was sufficient to achieve recognition as assayed in an appetitive conditioning paradigm [4]. Similarly, Rokni et al. showed that discriminability, or identification of a target odorant, decreased as similarity with the background increased. Difficulty of the task could be measured by the extent of overlap in glomerular activity; the greater the overlap, the more difficult the task [63]. Still, the system is able to adapt and identifying targets in varying background conditions.

Similarly, a study performed in *Drosophila melanogaster* has shown that responses of ORNs are invariant to changes in airspeed as long as stimulus concentrations are maintained [66]. Other systems, however, have been shown to manipulate changes in airflow to better extract information content. In both mouse and locust models, manipulation of airflow was shown to increase information about the odorant or source[67, 68].

Humidity is also a variable that influences detection or recognition of an odorant. Previous work understanding the role of relative humidity (RH) in olfaction has focused primarily on the odorant percept and identification. Studies conducted with both human subjects [69] and animal models report that higher humidity levels allow for lower concentrations of an odorant to be detected or that behavioral response is altered [70]. However, many others report no significant changes in behavioral performance under varying RH conditions [71, 72]. Detection of humidity is so important for insects that several species including, but not limited to, fruit flies (D. melanogaster) [73] and locusts [10, 17] have specialized sensors for its detection.

Biological olfaction has found numerous different solutions to remain invariant to environmental changes or to hijack properties of those changes to better extract information.

1.5.3 Sensor sensitivity and specificity

The final shortcoming of artificial olfaction I will discuss is in this section is achieving a high degree of specificity, or accurate identification of an odorant. There are many techniques used in electronic nose implementations with varying success, but as of yet nothing comes close to the versatility of biological olfactory systems in correctly identifying a large variety of odorants.

Natural olfactory systems have a number of features that enhance odorant discrimination. The first is a very large sensor array. Systems are made up of thousands of sensors, ORNs, which add a great deal of redundancy and improve signal to noise. For example, there are up to 1,200 ORNs in drosophila [74] and 50,000 ORNs in locusts [21, 23].

There is also a reorganization of the information as it proceeds through the pathway. Each ORN expresses only one type of olfactory receptor and all ORNs expressing the same receptors converge on the same groups of glomeruli. This convergence must allow for noise reduction [7, 8]. Since several PNs synapse to each glomeruli, and each PN can synapse into more than one glomerulus, they tend to have a broad response range. There is further reorganziation, as signals from the antennal lobe re-diverge in the KCs. In this reorganization, ~830 PNs are projected onto 50,000 KCs [21]. The divergence onto KCs as well as activation of inhibitory circuits in the mushroom body and lateral horn is hypothesized to create an action potential coincidence detector in which different KCs will respond to different odorants and concentrations [30].

Designers of electronic noses have attempted to replicate features of the olfactory system in order to better improve their sensors. To mimic the number of ORNs, a large array of sensors or pseudo-sensors is often utilized [47, 75-80]. However, fully replicating the scale of the sensor array in biological systems has of yet been infeasible and impractical. Furthermore, attempts to replicate the processing of the olfactory circuits have also been made. Techniques such as multi-layer perceptron (MLP), k-nearest-neighbor (kNN) classification, partial-least-squares discriminant analysis (PLS-DA), and dimensionality reduction, among other techniques have been attempted [81]. Despite numerous attempts, a true electronic nose still eludes proponents of the technology.

1.6 Thesis outline

In this thesis I will explore the three limitations of electronic noses and either propose solutions and alternatives to those challenges or elucidate the mechanisms in biological olfaction that overcome those issues. First, I will first detail all experimental and analytical methods in Chapter 2. Then in Chapter 3, I will describe a stimulation and analysis method to overcome sensor drift from both use and aging. Importantly, this technique allows for device replacement after sensor failure. Portions of this work have been published in Sensors and Actuators B: Chemical (2016). Chapter 4 will explore the effects of humidity modulation in a biological olfactory system. Here we gain insight in the mechanisms of humidity invariance as well as how changes in humidity are encoded in olfactory processing. In chapter 5, I will show analysis identifying malaria biomakers from human breath collected from children in Malawi. Finally, I will conclude this work in Chapter 6 explaining how it all integrates together as well as describe on going work developing a novel sensor that allows for high sensitivity and specificity to a target VOC used in malaria diagnosis.

1.7 References

1. Reymond, J.-L., *The Chemical Space Project.* Accounts of Chemical Research, 2015. **48**(3): p. 722-730.

2. Hildebrand, J.G. and G.M. Shepherd, *MECHANISMS OF OLFACTORY DISCRIMINATION:Converging Evidence for Common Principles Across Phyla.* Annual Review of Neuroscience, 1997. **20**(1): p. 595-631.

3. Simões, P., S.R. Ott, and J.E. Niven, *Associative olfactory learning in the desert locust, Schistocerca gregaria.* The Journal of Experimental Biology, 2011. **214**(15): p. 2495-2503.

4. Saha, D., et al., *A spatiotemporal coding mechanism for background-invariant odor recognition.* Nat Neurosci, 2013. **16**(12): p. 1830-1839.

5. Raman, B., et al., *Temporally Diverse Firing Patterns in Olfactory Receptor Neurons Underlie Spatiotemporal Neural Codes for Odors.* The Journal of Neuroscience, 2010. **30**(6): p. 1994-2006.

6. Boeckh, J., K.E. Kaissling, and D. Schneider, *Insect Olfactory Receptors.* Cold Spring Harbor Symposia on Quantitative Biology, 1965. **30**: p. 263-280.

7. Buck, L. and R. Axel, *A novel multigene family may encode odorant receptors: A molecular basis for odor recognition.* Cell, 1991. **65**(1): p. 175-187.

8. Saha, D. and B. Raman, *Relating early olfactory processing with behavior: a perspective.* Current Opinion in Insect Science, 2015. **12**: p. 54-63.

9. Kay, L.M. and M. Stopfer, *Information processing in the olfactory systems of insects and vertebrates.* Seminars in Cell & Developmental Biology, 2006. **17**(4): p. 433-442.

10. Ochieng', S.A., E. Hallberg, and B.S. Hansson, *Fine structure and distribution of antennal sensilla of the desert locust, Schistocerca gregaria (Orthoptera: Acrididae).* Cell and Tissue Research, 1998. **291**(3): p. 525-536.

11. Ameismeier, F., *Ultrastructure of the chemosensitive basiconic single-walled wall-pore sensilla on the antennae in adults and embryonic stages of Locusta migratoria L. (Insecta, Orthoptera).* Cell and Tissue Research, 1987. **247**(3): p. 605-612.

12. Ochieng', S.A. and B.S. Hansson, *Responses of olfactory receptor neurones to behaviourally important odours in gregarious and solitarious desert locust,Schistocerca gregaria.* Physiological Entomology, 1999. **24**(1): p. 28-36.

13. Greenwood, M. and R.F. Chapman, *Differences in numbers of sensilla on the antennae of solitarious and gregarious Locusta migratoria L. (Orthoptera: Acrididae).* International Journal of Insect Morphology and Embryology, 1984. **13**(4): p. 295-301.

14. Chapman, R.F. and M. Greenwood, *Changes in distribution and abundance of antennal sensilla during growth of Locusta migratoria L. (Orthoptera : Acrididae).* International Journal of Insect Morphology and Embryology, 1986. **15**(1): p. 83-96.

15. Hansson, B.S., et al., *Physiological responses and central nervous projections of antennal olfactory receptor neurons in the adult desert locust, Schistocerca gregaria (Orthoptera: Acrididae).* Journal of Comparative Physiology A, 1996. **179**(2): p. 157-167.

16. Steinbrecht, R., *Comparative morphology of olfactory receptors.* Olfaction and taste, 1969. **3**: p. 3-21.

17. Altner, H., C. Routil, and R. Loftus, *The structure of bimodal chemo-, thermo-, and hygroreceptive sensilla on the antenna of Locusta migratoria.* Cell and Tissue Research, 1981. **215**(2): p. 289-308.

18. Kafka, W.A., *Molekulare Wechselwirkungen bei der Erregung einzelner Riechzellen.* Zeitschrift für vergleichende Physiologie, 1970. **70**(2): p. 105-143.

19. Boeckh, J., *Die Reaktionen olfaktorischer Neurone im Deutocerebrum von Insekten im Vergleich zu den Antwortmustern der Geruchssinneszellen.* Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 1974. **90**(2): p. 183-205.

20. Leitch, B. and G. Laurent, *GABAergic synapses in the antennal lobe and mushroom body of the locust olfactory system.* The Journal of Comparative Neurology, 1996. **372**(4): p. 487-514.

21. Laurent, G., *Dynamical representation of odors by oscillating and evolving neural assemblies.* Trends in Neurosciences, 1996. **19**(11): p. 489-496.

22. Laurent, G., et al., *ODOR ENCODING AS AN ACTIVE, DYNAMICAL PROCESS: Experiments, Computation, and Theory.* Annual Review of Neuroscience, 2001. **24**(1): p. 263-297.

23. Stern, M., et al., *Regeneration of olfactory afferent axons in the locust brain.* The Journal of Comparative Neurology, 2012. **520**(4): p. 679-693.

24. Ernst, K.D., J. Boeckh, and V. Boeckh, *A neuroanatomical study on the organization of the central antennal pathways in insects.* Cell and Tissue Research, 1977. **176**(3): p. 285-308.

25. Wang, F., et al., *Odorant Receptors Govern the Formation of a Precise Topographic Map.* Cell, 1998. **93**(1): p. 47-60.

26. Vosshall, L.B., A.M. Wong, and R. Axel, *An Olfactory Sensory Map in the Fly Brain.* Cell, 2000. **102**(2): p. 147-159.

27. MacLeod, K. and G. Laurent, *Distinct Mechanisms for Synchronization and Temporal Patterning of Odor-Encoding Neural Assemblies.* Science, 1996. **274**(5289): p. 976-979.

28. Laurent, G., *Olfactory network dynamics and the coding of multidimensional signals.* Nat Rev Neurosci, 2002. **3**(11): p. 884-895.

29. Stopfer, M., V. Jayaraman, and G. Laurent, *Intensity versus Identity Coding in an Olfactory System.* Neuron, 2003. **39**(6): p. 991-1004.

30. Stopfer, M., *Central processing in the mushroom bodies.* Current Opinion in Insect Science, 2014. **6**: p. 99-103.

31. Papadopoulou, M., et al., *Normalization for Sparse Encoding of Odors by a Wide-Field Interneuron.* Science, 2011. **332**(6030): p. 721-725.

32. Martin, A.J.P. and R.L.M. Synge, *A new form of chromatogram employing two liquid phases.* Biochemical Journal, 1941. **35**(12): p. 1358-1368.

33. James, A.T. and A.J.P. Martin, *Gas-liquid partition chromatography: the separation and micro-estimation of volatile fatty acids from formic acid to dodecanoic acid.* Biochemical Journal, 1952. **50**(5): p. 679-690.

34. Rose, B.A., *Gas chromatography and its analytical applications. A review.* Analyst, 1959. **84**(1003): p. 574b-595.

35. Bartle, K.D. and P. Myers, *History of gas chromatography.* TrAC Trends in Analytical Chemistry, 2002. **21**(9–10): p. 547-557.

36. Grob, R.L. and E.F. Barry, *Modern practice of gas chromatography*. 2004, Hoboken, N.J.: Wiley-Interscience.

37. Toko, K., *Biomimetic Sensor Technology*, 2000, Cambridge University Press.

38. Seiyama, T., et al., *A New Detector for Gaseous Components Using Semiconductive Thin Films.* Analytical Chemistry, 1962. **34**(11): p. 1502-1503.

39. Azad, A.M., et al., *Solid-State Gas Sensors: A Review.* Journal of The Electrochemical Society, 1992. **139**(12): p. 3690-3704.

40. Göpel, W. and K.D. Schierbaum, *SnO2 sensors: current status and future prospects.* Sensors and Actuators B: Chemical, 1995. **26**(1): p. 1-12.

41. Moseley, P.T.T.B.C., *Solid-state gas sensors*. 1987, Bristol; Philadelphia: A. Hilger.

42. Semancik, S., et al., *Microhotplate platforms for chemical sensor research.* Sensors and Actuators B: Chemical, 2001. **77**(1–2): p. 579-591.

43. Clifford, P.K. and D.T. Tuma, *Characteristics of semiconductor gas sensors II. transient response to temperature change.* Sensors and Actuators, 1982. **3**(0): p. 255-281.

44. Afridi, M.Y., et al., *A monolithic CMOS microhotplate-based gas sensor system.* IEEE Sensors Journal, 2002. **2**(6): p. 644-655.

45. Williams, D.E., *Semiconducting oxides as gas-sensitive resistors.* Sensors and Actuators B: Chemical, 1999. **57**(1–3): p. 1-16.

46. Hierlemann, A. and R. Gutierrez-Osuna, *Higher-Order Chemical Sensing.* Chemical Reviews, 2008. **108**(2): p. 563-613.

47. Rogers, P.H. and S. Semancik, *Development of optimization procedures for application-specific chemical sensing.* Sensors and Actuators B: Chemical, 2012. **163**(1): p. 8-19.

48. Artursson, T., et al., *Drift correction for gas sensors using multivariate methods.* Journal of Chemometrics, 2000. **14**(5-6): p. 711-723.

49. Raman, B., et al., *Bioinspired Methodology for Artificial Olfaction.* Analytical Chemistry, 2008. **80**(22): p. 8364-8371.

50. Raman, B., et al., *Detecting and recognizing chemical targets in untrained backgrounds with temperature programmed sensors.* Sensors Journal, IEEE, 2012. **12**(11): p. 3238-3247.

51. Vergara, A., et al., *Chemical gas sensor drift compensation using classifier ensembles.* Sensors and Actuators B: Chemical, 2012. **166–167**(0): p. 320-329.

52. Zuppa, M., et al., *Drift counteraction with multiple self-organising maps for an electronic nose.* Sensors and Actuators B: Chemical, 2004. **98**(2–3): p. 305-317.

53. Holmberg, M., et al., *Drift counteraction in odour recognition applications: lifelong calibration method.* Sensors and Actuators B: Chemical, 1997. **42**(3): p. 185-194.

54. Saha, D., et al., *Behavioural correlates of combinatorial versus temporal features of odour codes.* Nat Commun, 2015. **6**.

55. Pearce, T.C., *Handbook of machine olfaction : electronic nose technology*. 2003, Weinheim Germany: Wiley-VCH. xxxii, 592 p.

56. Graziadei, P.P.C. and G.A.M. Graziadei, *Neurogenesis and neuron regeneration in the olfactory system of mammals. I. Morphological aspects of differentiation and structural organization of the olfactory sensory neurons.* Journal of Neurocytology, 1979. **8**(1): p. 1-18.

57. Graziadei, P.P.C. and G.A.M. Graziadei, *Neurogenesis and Plasticity of the Olfactory Sensory Neuronsa.* Annals of the New York Academy of Sciences, 1985. **457**(1): p. 127-142.

58. Hinds, J.W. and P.L. Hinds, *Synapse formation in the mouse olfactory bulb. II. Morphogenesis.* The Journal of Comparative Neurology, 1976. **169**(1): p. 41-61.

59. Hinds, J.W. and P.L. Hinds, *Synapse formation in the mouse olfactory bulb Quantitative studies.* The Journal of Comparative Neurology, 1976. **169**(1): p. 15-40.

60. Riecke, H., *Olfactory Computation and Adult Neurogenesis.* 2013.

61. Gheusi, G. and P.-M. Lledo, *Control of Early Events in Olfactory Processing by Adult Neurogenesis.* Chemical Senses, 2007. **32**(4): p. 397-409.

62. Saha, D., et al., *Multi-unit Recording Methods to Characterize Neural Activity in the Locust (Schistocerca Americana) Olfactory Circuits.* 2013(71): p. e50139.

63. Rokni, D., et al., *An olfactory cocktail party: figure-ground segregation of odorants in rodents.* Nat Neurosci, 2014. **17**(9): p. 1225-1232.

64. Laing, D.G., H. Panhuber, and B.M. Slotnick, *Odor masking in the rat.* Physiology & Behavior, 1989. **45**(4): p. 689-694.

65. Laing, D.G. and G.W. Francis, *The capacity of humans to identify odors in mixtures.* Physiology & Behavior, 1989. **46**(5): p. 809-814.

66. Zhou, Y. and R.I. Wilson, *Transduction in Drosophilaolfactory receptor neurons is invariant to air speed.* Journal of Neurophysiology, 2012. **108**(7): p. 2051-2059.

67. Oka, Y., Y. Takai, and K. Touhara, *Nasal Airflow Rate Affects the Sensitivity and Pattern of Glomerular Odorant Responses in the Mouse Olfactory Bulb.* The Journal of Neuroscience, 2009. **29**(39): p. 12070-12078.

68. Huston, Stephen J., et al., *Neural Encoding of Odors during Active Sampling and in Turbulent Plumes.* Neuron, 2015. **88**(2): p. 403-418.

69. Kuehn, M., et al., *Changes of pressure and humidity affect olfactory function.* European Archives of Oto-Rhino-Laryngology, 2008. **265**(3): p. 299-302.

70. Krång, A.-S., et al., *Transition from sea to land: olfactory function and constraints in the terrestrial hermit crab Coenobita clypeatus.* Proceedings of the Royal Society B: Biological Sciences, 2012. **279**(1742): p. 3510-3519.

71. Philpott, C., et al., *The effect of temperature, humidity and peak inspiratory nasal flow on olfactory thresholds.* Clinical Otolaryngology & Allied Sciences, 2004. **29**(1): p. 24-31.

72. Alfonso Collado, J.I. and H. Vallés Varela, *Study of Olfactory Function for Pyridine in Healthy Population: Influence of Variations in Humidity.* Acta Otorrinolaringologica (English Edition), 2008. **59**(10): p. 475-479.

73. Enjin, A., et al., *Humidity Sensing in Drosophila.* Current Biology, 2016. **26**(10): p. 1352-1358.

74. Stocker, R.F., *The organization of the chemosensory system in Drosophila melanogaster: a rewiew.* Cell and Tissue Research, 1994. **275**(1): p. 3-26.

75. Cavicchi, R.E., et al., *Optimized temperature-pulse sequences for the enhancement of chemically specific response patterns from micro-hotplate gas sensors.* Sensors and Actuators B: Chemical, 1996. **33**(1–3): p. 142-146.

76. Kunt, T.A., et al., *Optimization of temperature programmed sensing for gas identification using micro-hotplate sensors.* Sensors and Actuators B: Chemical, 1998. **53**(1–2): p. 24-43.

77. Boger, Z., et al., *Rapid Identification of Chemical Warfare Agents by Artificial Neural Network Pruning of Temperature-Programmed Microsensor Databases.* Sensor Letters, 2003. **1**(1): p. 86-92.

78. Vergara, A., et al., *Optimized temperature modulation of micro-hotplate gas sensors through pseudorandom binary sequences.* Sensors Journal, IEEE, 2005. **5**(6): p. 1369-1378.

79. Benkstein, K., et al., *Integration of nanostructured materials with MEMS microhotplate platforms to enhance chemical sensor performance.* Journal of Nanoparticle Research, 2006. **8**(6): p. 809-822.

80. Vergara, A., et al., *Quantitative gas mixture analysis using temperature-modulated micro-hotplate gas sensors: Selection and validation of the optimal modulating frequencies.* Sensors and Actuators B: Chemical, 2007. **123**(2): p. 1002-1016.

81. Gutierrez-Osuna, R., *Pattern analysis for machine olfaction: a review.* Sensors Journal, IEEE, 2002. **2**(3): p. 189-202.

Chapter 2: Methods

2.1 Sensor testing and development

2.1.1 Development of a fully automated analyte delivery system

Precise generation and delivery of gases and their mixtures at relevant concentrations is critical for our chemical sensor studies. A schematic of the designed and validated odor delivery system is shown in Figure 2.1. The current gas generation system has been designed to fulfill the following goals: (i) continuous generation of gas mixtures of known composition up to 24 hours; (ii) generation of a wide range of gas flow from 10 mL/min to 2000 mL/min; (iii) wide dynamic range (from parts-per-billion to percentage levels); (iv) dynamic and real-time control over the gas concentration and the flow rate across the sensor array, and (v) fully automated chemical delivery setup where flow rate, concentration and exposure time can be controlled via a graphical user interface software.

Our odor delivery system has five input flow lines and can generate complex mixtures of various concentrations which can be delivered at varying flow rates. The gas sources are from either a gas cylinder or a gas bubbler. Gas bubblers are bottles partially filled with an odorant through which clean air is bubbled. This maintains a constantly saturated headspace in the remaining volume of the bubbler. This saturated headspace is injected into the system to create a high-concentration gas streams, which are diluted by separate background flow and dilutions lines. Maximum flowrates from each mass flow controller are indicated in Fig. 2.1. Lower concentration (sub-ppb range) can be obtained by using a second or even a third dilution step before injecting the flow into the sensing manifold. This flow system was used for all sensor testing and validation experiments.

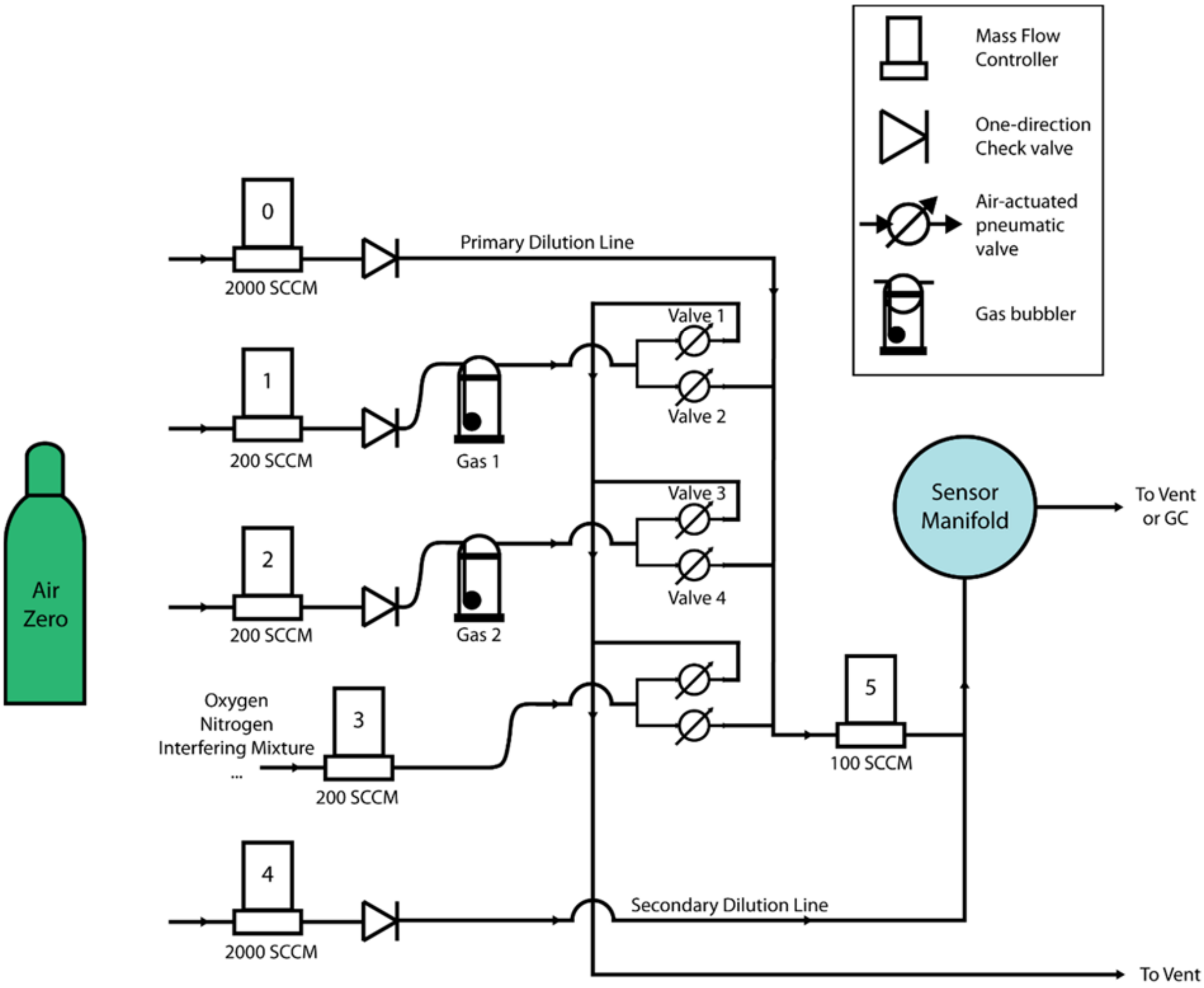


Figure 2.1. Analyte-delivery system for chemical sensing experiments. Flow system allows for precise control of gas concentration, mixture, and flow rate for testing and development of gas sensors. Flow rates to the sensing manifold range from 0.01 to 2.1 SLM. Concentrations of chemicals delivered span from the high PPB range to percentiles.

2.1.2 Analyte delivery

Consistent I/O transforms

Liquid chemicals were vaporized in 500 mL gas washing bottles (Pyrex, Corning, NY) using zero air (Airgas, St. Louis, MO). 0.1 standard l/min (slm) of this saturated vapor was diluted in a 1 slm stream of filtered, dehumidified air. Only 100 standard cm^3/min (sccm) of this diluted stream mixed with 650 sccm of filtered, dehumidified air was finally directed to the sensor manifold. Regardless of whether an analyte was present, a constant air flow of 750 sccm across the sensor was maintained at all times. The analyte for each trial was selected pseudo-randomly to reduce effects of long-term chemical hysteresis. The analytes were presented in a pulsed fashion as shown in Figure 2.2.

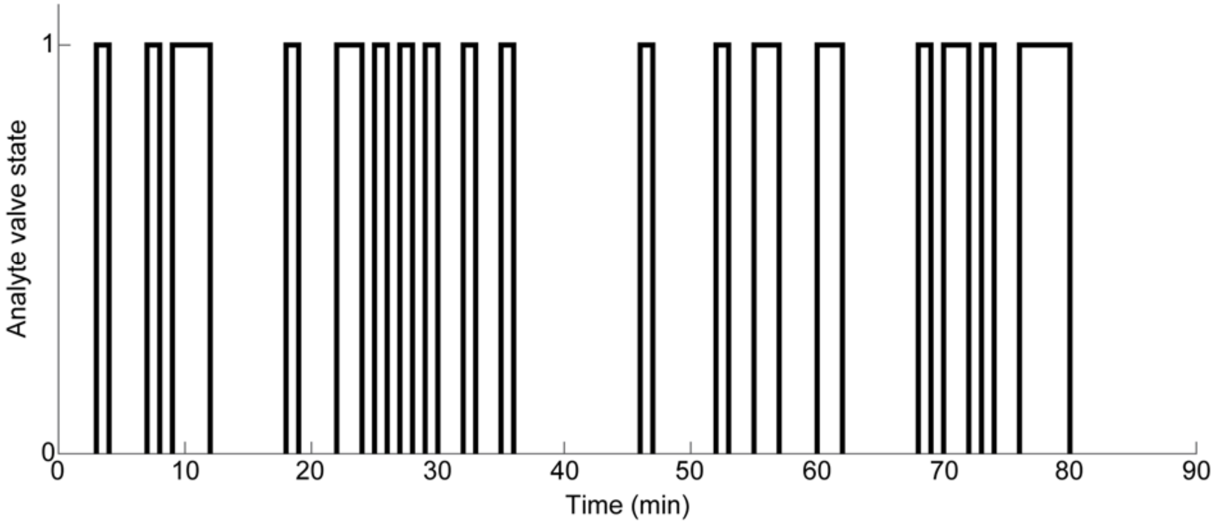


Figure 2.2. M-sequence stimulation profile. Valve state of 0 indicates analyte gas is vented. Valve state of 1 indicates analyte gas is flowing to sensor manifold.

For altering analyte concentrations, we varied the relative volume of the saturated output from the bubblers and the first carrier stream, before sub-sampling (i.e. during the first dilution stage). Note that only the volume of the first carrier stream was changed; the flowrate through the bubbler was kept constant. The total flow rate of the first carrier stream was varied from a 1 slm flow rate to 0.5, 0.75, or 1.25 slm. The resulting analyte concentrations achieved as a result of this flow modulation are shown in Figure 2.3.

The following five chemicals were used in this study: ethanol (Pharmco-Aaper, Brookfield, CT), acetone, hexanol, 2-octanol, and 2-octanone (Sigma-Aldrich, St. Louis, MO).

	Analyte Concentration (µmol/mol)			
Acetone	6.7E+03	4.7E+03	3.6E+03	3.0E+03
Ethanol	1.72E+03	1.21E+03	9.36E+02	7.63E+02
Hexanol	1.918E+01	1.354E+01	1.046E+01	8.526E+00
2-Octanol	7.0E+00	4.9E+00	3.8E+00	3.1E+00
2-Octanone	3.2E+01	2.3E+01	1.8E+01	1.4E+01
Dilution Line Flowrate (SLM)	0.5	0.75	1	1.25

Figure 2.3. Analyte concentration as a function of dilution line flow rate. Concentration differences arise from differences in analyte vapor pressure; note that flow rates were kept consistent across analytes. Table and legend as is from [1].

Crumpled graphene oxide chemiresistive sensors

Liquid chemicals were vaporized in customized airfree bubblers (Chemglass Life Sciences, Vineland, NJ) using zero air (Airgas, St. Louis, MO). Flow rates through the system varied and depended on the concentration to be delivered. Carrier airstreams were filtered, dehumidified compressed air. Total airflow rate to the sensing manifold was a constant 750 sccm, regardless of whether an analyte was present. The analyte for each trial was selected pseudo-randomly. The analytes were presented in 5 minute pulses with 5 minutes between sequential pulses.

The following 4 chemicals were used in this study: 1-octanol, 2-octanol, geraniol, and α-pinene (Sigma-Aldrich, St. Louis, MO).

2.1.3 Metal oxide sensor

A microsensor array with four individually controllable elements covered by SnO2 sensing films was used in this study (Figure 2.4). Other sensing elements in the array were left idle. The manufacture of these devices has been thoroughly described previously [2-7]. Briefly, each sensor element is a multilayer, suspended device. From the top, the functional layers are: a polycrystalline SnO2 sensing film, two interdigitated platinum electrodes, an insulating layer, and a polysilicon heater. The operating temperature of the sensor was modulated between 55 °C and 435 °C. Sensor responses of four copies of the SnO2 microsensors were measured. Each sensor was cycled through 28 temperature steps (Figure 2.5), with each temperature treated as a perturbed-isotherm [8].

Previous work by us [9], has shown that at least two different correlated bands of information were generated at low and high temperatures for most analytes. Therefore, we used a temperature program that sampled a range of temperatures between 55 °C – 435 °C.

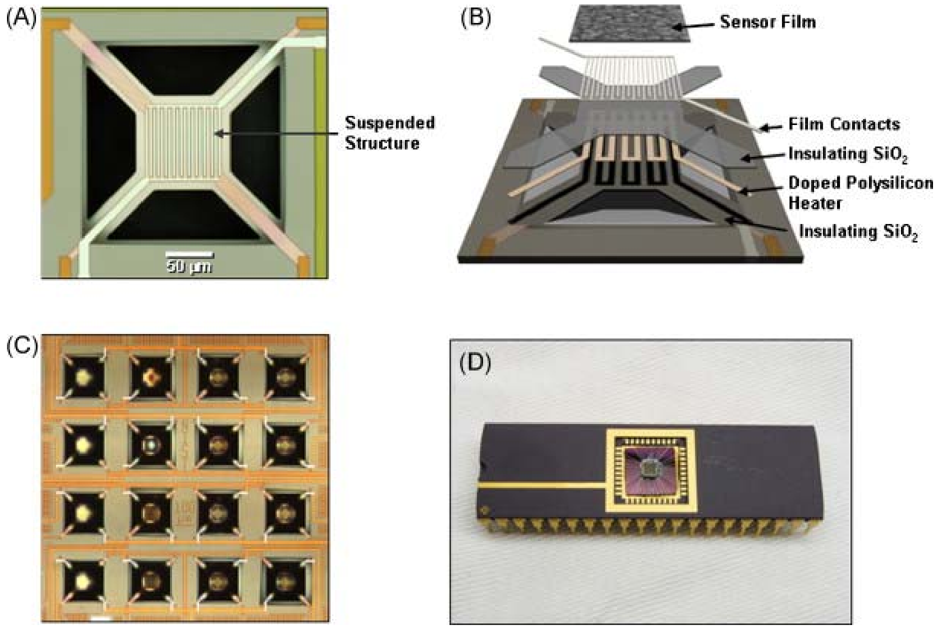


Figure 2.4. Micro-hotplate sensor array platform. (A) An optical microscopy image of a single microhotplate microsensor element. (B) A layered schematic showing the three primary components of the microsensor elements: polycrystalline silicon heater, interdigitated platinum electrodes, and metal oxide sensing film. (C) A microsensor array with 16 individually addressable, temperature-controlled elements. (D) A 40-pin dual in-line packaged microsensor device. Figure as is from [9].

Sensor responses of four copies of the SnO2 microsensors were measured. Each sensor was cycled through 28 temperature steps (Figure 2.5), with each temperature treated as a perturbed-isotherm [8] for that particular microsensor. Conductance measurements were made at each of the operation steps for each sensor. Each measurement cycle lasted 38 s as all four sensors used were cycled through 28 temperatures. A logarithm (base 10) was calculated to compress the sensor responses (note that this step is not critical for results reported here). Analysis in this paper was performed using three sensors that lasted the entire data collection period of approximately 7 months.

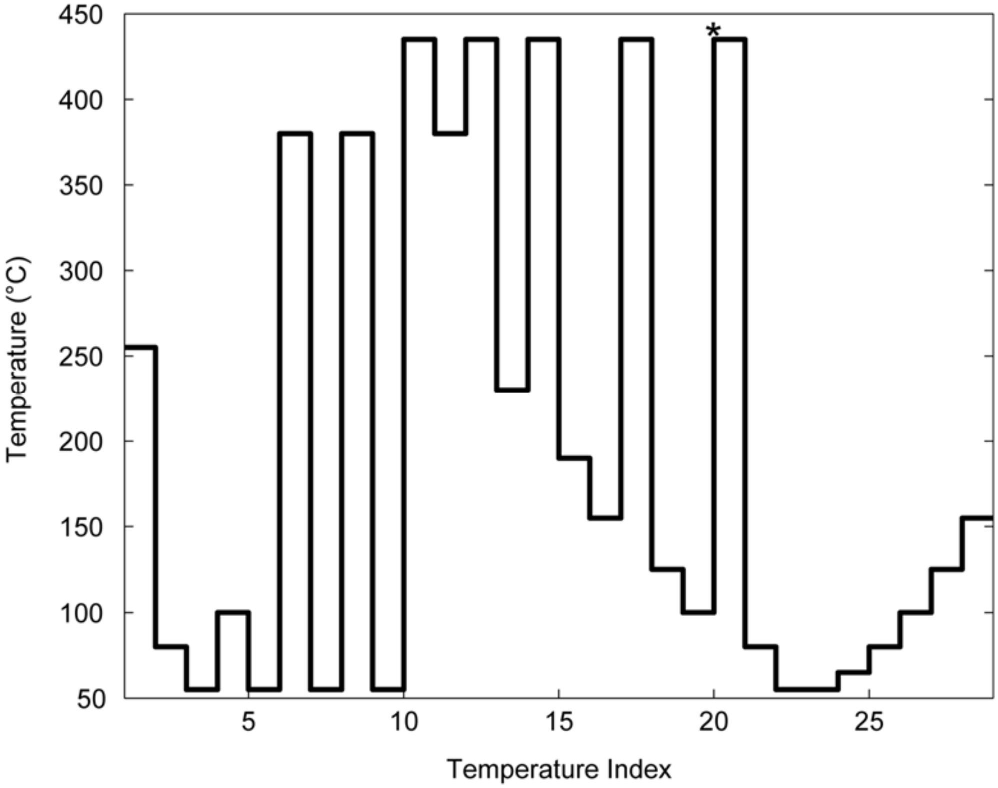


Figure 2.5. Temperature cycle employed for micro-hotplates. A data point was collected from each sensor, at these temperatures, for every time point. The sensor was allowed to dwell at a temperature for 750 ms before a measurement was taken and each temperature was treated as a single trace. The star marks the 20th temperature index on which most of the analysis was performed. Figure and legend as is from [1].

2.1.4 Calculation of the I/O transforms

The sensor response (y) at a particular temperature was treated as the dependent variable to be predicted given the most recent stimulus history ($\vec{x}$; series of 1 and 0 indicating analyte valve open or closed during each measurement). Note, the length of the moving window (i.e. dimensionality of vector ($\vec{x}$)) is a free parameter. For each sensor, temperature combination:

$$y = \vec{k} \cdot \vec{x} \qquad (2.1)$$

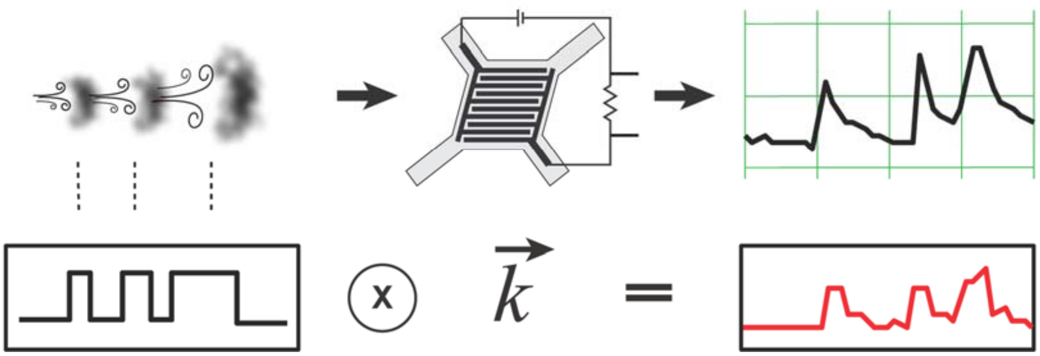


Figure 2.6. Schematic of I/O transform. A) Top row) Schematic illustration of the proposed approach. An analyte is pulsed over the sensor and causes changes in resistance across a metal oxide sensing film. Bottom row) A schematic of a sensor response reconstruction is shown. Given a specific stimulus sequence, the sensor's operation is akin to a mathematical transformation that is specific for a given analyte. Convolving the analyte-specific filter with the most recent stimulus history will generate a prediction for the sensor's response that can be expected at the end of that period.

where $\vec{k}$ is the transform that maps the stimulus history (the input) onto the sensor's response (the output) for a given analyte (Figure 2.6). I/O transform for each analyte was calculated using a least squares regression estimation method, where the mean squared error of prediction is defined as:

$$(\vec{k} \cdot \vec{x} - y)^2 \qquad (2.2)$$

A moving window was continuously shifted and the sensor response at the end of the stimulus history was recorded to construct a sensor input matrix (X), and a sensor response vector (Y; a column vector of sensor responses at the end of each stimulus history). A bias term was added to account for average signal baseline (DC offset) throughout the experiment. This bias-term was not used in any further analysis as the goal was to focus primarily on the changes in response dynamics. Note that the reconstruction error is minimized when:

$$\vec{k} = \left(\vec{X}^T \vec{X}\right)^{-1} X^T Y \qquad (2.3)$$

i.e. the pseudo-inverse solution.

For results shown in this manuscript, a stimulus history of eight recent stimulus states was chosen as the window size. This selection was based on the minimum window length for which the error of reconstruction converged (Figure 2.7). This was done to provide a sufficient tradeoff between minimizing reconstruction error and reducing over-fitting to the training data.

Note that the response of the sensor at each temperature was treated separately. Therefore, we created 28 I/O transforms for each analyte one for each operating temperature. These are shown in Fig. 3.2C.

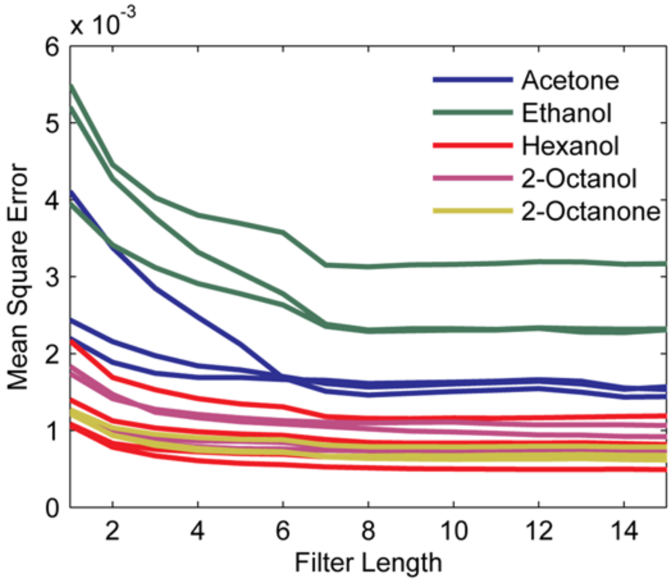


Figure 2.7. Mean square error of sensor response reconstruction as a function of filter length. Colors represent different analytes, each trace is a different run from the first dataset collected.

2.1.5 Training and testing datasets

First dataset: The sensor response measurements were collected over a period of a month. Each analyte was presented in a random binary pulsatile sequence (analyte ON and OFF) as shown in Fig. 1. A single pulsatile sequence run resulted in 140 sensor response measurements (~40 with analyte ON period and ~100 analyte OFF period). The training data set consisted of 17 such measurement sequences; 3 for acetone, 3 for ethanol, 4 for hexanol, 4 for 2-octanol, and 3 for 2-octanone. Note that this was the training data used for generating the I/O transforms shown in Fig 3.2, and 3.4.

Second dataset: The sensor array was purposely aged for a period of two months during which it was intermittently operated with exposure to the same analytes. Subsequent to the aging phase, a second experimental run, lasting a 2-3 weeks of data collection, was used to validate our approach. The second experimental run consisted of similar sequences as in the training period. 20 such measurement sequences were made (4 for each analyte). This data was used as the testing data for quantifying performance of the models generated using the first training dataset (refer Fig. 3.7).

Third dataset (Concentration dataset): To further assess the limits of our methodology a third dataset using a subset of the analytes (acetone, ethanol, hexanol, and 2-octanol) was presented at varying concentrations. Note that this third dataset was collected 3 months after the collection of the second dataset. This dataset consisted of 13 measurement sequences; 4 for Acetone, 3 for Ethanol, 3 for Hexanol, and 3 for 2-Octanol. Also note that the dataset collected during the previous data collection phase (i.e. second dataset) were used to create the I/O transforms for classifying these responses (Fig. 3.9).

2.1.6 Dimensionality reduction and classification.

A linear principle component analysis (PCA) was used for visualizing the I/O transforms associated with each analyte. The I/O transforms or response filters were projected onto eigenvectors corresponding to the three largest eigenvalues of the correlation matrix. Only the training data and their corresponding transforms were used for determining the PCA axes. Both training and testing datasets were projected onto the same axes to aid visualization and comparison.

In order to classify testing data, we implemented a hierarchical, recursive approach. In this scheme, we treated one m-class classification problem into a series of m binary classification problem. The algorithm is shown in schematic form in Figure 2.8 and can be summarized as follows:

1. Compute the I/O transform of the test analyte that needs to be classified
2. Perform PCA to reduce dimensionality of the filter (Note that this step was performed for visualization, but not strictly necessary for the approach).
3. Compute the mean centroid of the dimensionally reduced I/O transform of each training analyte
4. Repeat until done:
    a. Based on pairwise distances between the mean I/O transforms, find the analyte that is farthest from all others (i.e. the analyte with the most unique I/O filter shape; in Fig. 5 this is ethanol).
    b. Project all training and test data onto a difference of mean plane between the centroid of the farthest analyte and the centroid of the closest analyte (for ethanol the closest response cluster is acetone).
    c. Use a nearest neighbor classifier to assign the test sample to one of the two classes: farthest analyte vs. rest of training
    d. If assigned to farthest analyte, then stop recursion.
    e. If only two training analytes remain, then classify the test sample using the label of the nearest neighbor in the training sample.

f. Else, remove the samples from the farthest analyte from further consideration and repeat steps a-e.

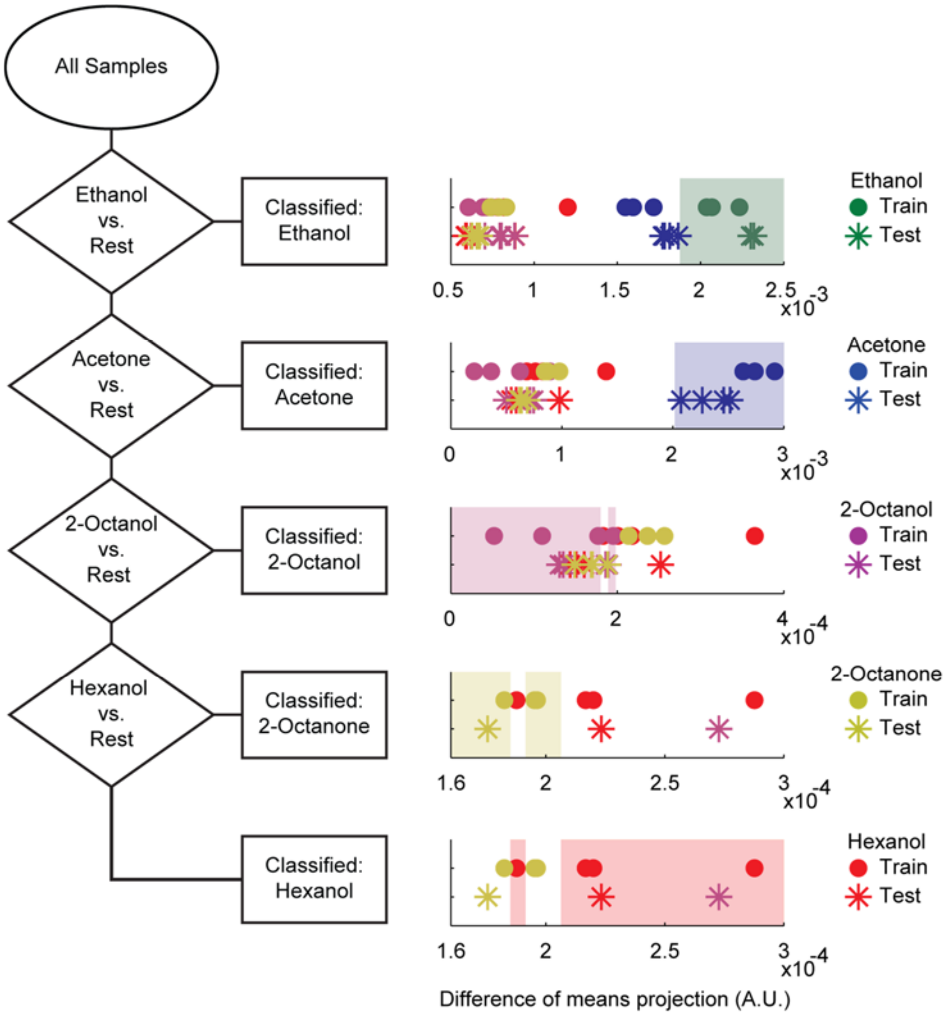


Figure 2.8. A hierarchical classification algorithm for multi-analyte recognition. Hierarchical classification algorithm used to identify analytes is schematically shown. At each level, both training and testing data were projected onto the differences of means plane between the most distinct/farthest class and its nearest neighboring class. The data after this projection is shown in the panels on the right. The regions where the class assignment favored the distant class are identified in each subplot displaying projected data. When testing data projects onto un-colored regions, those points will move down a level in order to be precisely classified.

2.1.7 Room temperature sensor fabrication

Electrode array fabrication

Interdigitated electrodes were made using standard micro-fabrication techniques. A top-down fabrication technique was employed, in which a metal layer is deposited onto a substrate (glass or Si wafer) and then selectively etched away. Deposition was done in a two-step process using thermal evaporation. First, the substrate was cleaned and a thin adhesion layer (~5-10 nm Cr) was applied to enable gold deposition. Next, a uniform layer of gold, approximately 50-75 nm, was deposited as the primary electrode material. Gold was selected as the electrode metal because of its high electrical conductivity, surface modification potential, and resistance to corrosion. Following metal deposition, photoresist was patterned on in the shape of the sensor array using standard photolithography methods. The general steps are outlined in Figure 2.9A.

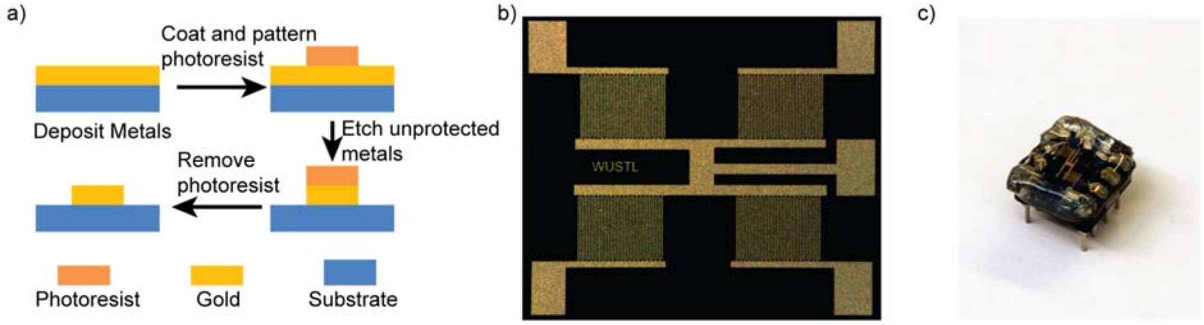


Figure 2.9. Fabrication of multi-electrode arrays for developing room temperature chemical sensing array. a) Top down nano-fabrication methodology to make sensor arrays. Gold is adhered to a substrate, a photoresist is spincoated and then patterned on top to protect gold in desired locations. Excess or undesired gold is removed or etched away, leaving only a desired pattern. b) A fully fabricated gold interdigitated electrode array. Electrode digit width is 20um, with 20 um between digits. c) Final, device with electrode array bonded and wired to DIP socket.

In a top-down technique, a uniform, 450 nm thick layer of photoresist (S1805, Shipley Co.), was spin coated onto a metal deposited substrate. The coated substrate was 'pre-baked' on a hotplate at 115C for 60 seconds and allowed to cool to room temperature. Then, in a mask aligner, a custom-made negative photomask was used to expose certain portions of the substrate to UV light. The substrate was then 'post-baked' on a hotplate at 115 °C for 60 seconds and allowed to cool to room temperature. Following the post-bake, the photoresist was developed for 40 seconds in MF-319 developer and the pattern shape was removed from the resist layer.

This photoresist layer protected the underlying gold during a liquid etching process using gold etchant TFA (Transene, Inc.). Unprotected gold on the substrate was removed. Next, reactive ion etching was used to remove the exposed Cr adhesion layer and any residual gold and photoresist along the edges. RIE etching was done using an Oxford Plasmalab 100 ICP-Reactive Ion Etcher. The recipe used consisted of a 15 SCCM flow of Cl2 and 5 SCCM flow of Ar in a chamber with 2 mT pressure and at 20C. The ICP power was set to 600W and the RF power was set to 250W. The standard etching time was 30 seconds, but time was increased or decreased as required based on the thickness of the deposited Au and Cr layers. Pre- and post-etching purges were done with 50 SCCM Ar for 2 minutes to remove any potential contaminants or hazards from the chamber. Finally, the photoresist was removed in Remover PG (MicroChem Corp.), leaving only the interdigitated electrode arrays (Figure 2.9B).

Chemiresistive film deposition

Prior to film deposition, sensor array was heated and maintained at 50 °C on a hotplate. Crumpled graphene oxide was dropcast on electrode in 0.5 – 2 µl drops, while being careful to restrict drops to a single electrode. Drops were repeatedly deposited until resistance across dry film was ~1MΩ.

Socket assembly

The electrode arrays were mounted to 8-pin DIP sockets (ED90032-ND, Digikey) using a small amount of dental wax (6652151, Henry-Schein Dental). Silver wires were pre-soldered to the dip socket, then connected to electrode pads with silver print (842, MG Chemicals, Surrey, B.C.). The electrode array and wires were permanently sealed into place by apply epoxy (84101, Permatex) to the outside edges. A fully assembled sensor is shown in Figure 2.9C.

2.1.8 Crumpled graphene oxide

Crumpled graphene oxide (cGO) was made following previously published methods [10]. In brief, an aqueous suspension of graphene oxide was aerosolized into 2-4 µm drops and heated in a furnace aerosol reactor at a minimum temperature of 400 °C. Compression due to capillary forces compressed the sheets into a crumpled 3D structure. The morphology of the sheets can be seen in Figure 2.10.

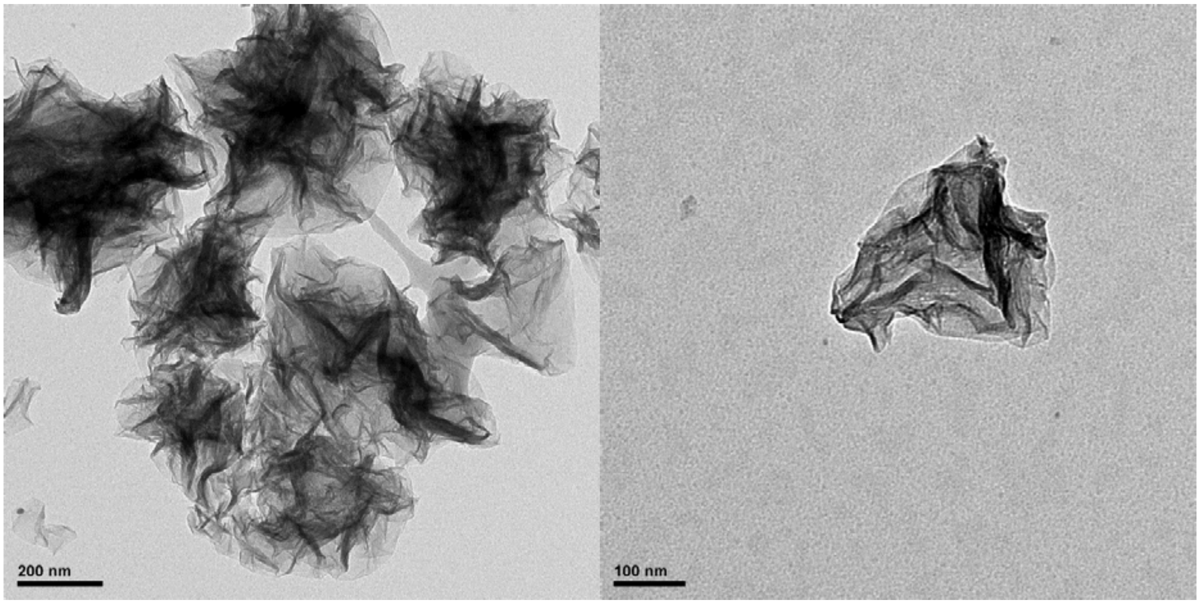


Figure 2.10. TEM images of crumpled graphene oxide.

2.2 Locust electrophysiology

2.2.1 Odor bottle preparation

The following five chemicals were used in this study: isoamyl acetate, hexanol, 1-octanol, and cyclohexanone (Sigma-Aldrich, St. Louis, MO). Analytes were diluted in paraffin oil to a 1% concentration (v/v) and 20 ml were placed in 60 ml glass bottle. Hexanol was also tested at 0.1% and 10% to compare stimulus intensities.

PREDICTED RH (%)	MEASURED RH (%)	TEMPERATURE (°C)
0	12.1	22.6
33	37.5	22.2
66	67.5	22.0
100	99.9	22.1
0	9.3	22.7

Figure 2.11. Hygrometer measurements of carrier stream humidity levels.

2.2.2 Odorant delivery and humidity control

Analytes were delivered following a standard protocol used in other works [11-14]. Analytes were delivered by injecting a constant volume (100 sccm) of the static headspace of the odor bottles into a carrier stream of filtered, dehumidified air (750 sccm) flowing continuously to the locust. To add humidity, a fraction of the carrier stream was diverted through a water bubbler to be humidified before being reintroduced with the remainder stream (Figure 2.12A). Flow rates of 0, 250, 500, or 750 sccm were diverted for humidification allowing for control of humidity levels at 0%, 33%, 66%, and 100% relative humidity of the carrier stream. Humidity levels were validated using traceable hygrometer (Model # 1166118, Fisher Scientific) and actual RH levels were close to expected levels (Figure 2.11). A large vacuum was placed behind the preparation to remove delivered odorants.

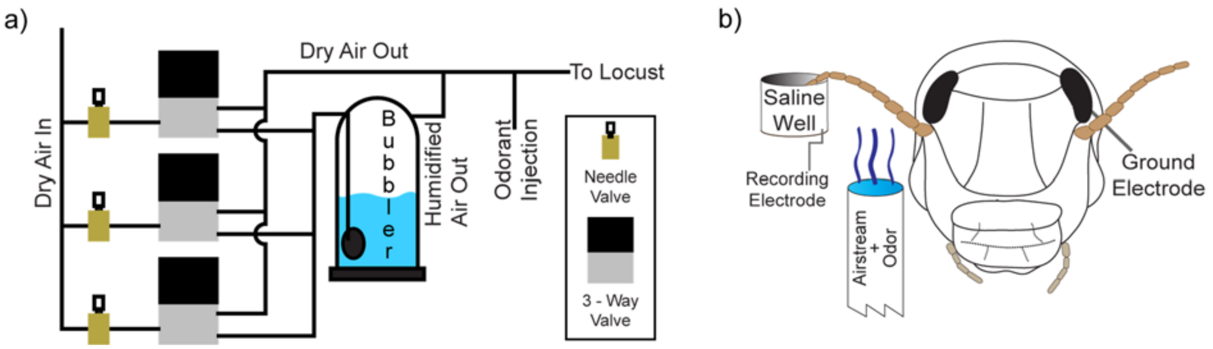


Figure 2.12. Humidification apparatus and Locust EAGs. a) Clean, dehumidified air flows through a valve system. Valves deliver air directly to the locust or through a bubbler for humidification before mixing with dry air delivered to the locust. The ratio of dry air vs humidified air controls overall humidity level of the airstream to the locust. b) Schematic drawing of locust EAG. Dorsal tip of an antenna is cut and dipped in a well of Locust Ringer's solution. A recording electrode is inserted in the well and ground electrode is inserted in the contra-lateral eye.

2.2.3 Locust electroantennogram (EAG)

EAGS have been subject to extraneous noise from environmental variability such as airflow changes. Furthermore, EAGs are often performed on excised antennae which slowly die after amputation [15]. This yields unreliable results over time. To overcome these limitations, we employed a novel EAG method that allowed for use of an intact locust prep, yet isolated the antenna from external environmental variations.

Recordings were performed on an intact locust (*Schistocerca americana*) preparation. First, the locust was restrained in a plastic tube which was inserted into a custom made 3D-printed chamber (Figure 2.13). One antenna was extended and the distal end was dipped in a well filled with locust ringer's solution [16]. The distal most segment was removed to expose antennal lymph and allow for electrical conductivity. A silver chloride wire was inserted into the saline, for the recording electrode, and another in the contra-lateral eye as a ground electrode. Wax was applied around the head to restrict movement and secure the ground wire (Figure

2.12B, 2.13B). Data was collected at 15 kHz with a custom made LabVIEW data acquisition program.

EAGS were collected from 10 locusts for 3 intensities of hexanol (0.1%, 1%, and 10%) under 4 humidity levels (0%, 33%, 66%, and 100% RH). Five trials for each intensity at each humidity level were collected. The order of intensity was pseudo-random for each experiment, but the order for humidity changes was fixed: 0%, 33%, 66%, and 100% before returning to 0%. A 5 minute acclimation period followed every switch from 100% to 0% RH.

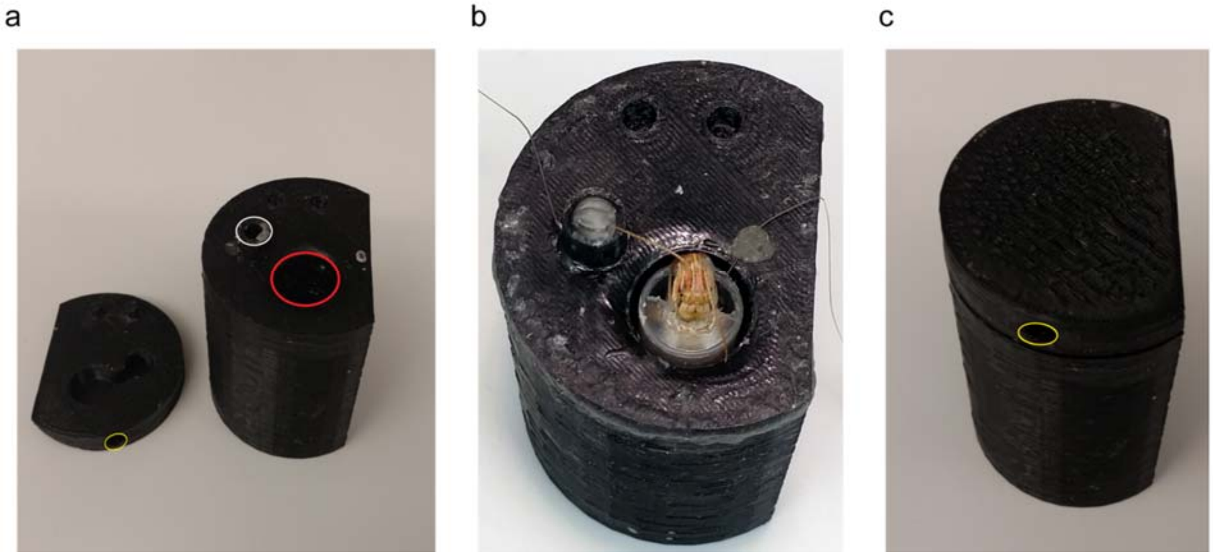


Figure 2.13. Custom designed manifold for locust electroantennogram recordings. a) Image of open chamber. Circles indicate important ports or openings. Yellow circle is airflow inlet, white circle is saline well, and red circle is opening for locust. Note that there is an airflow outlet is present opposite the inlet, so air may flow through unobstructed. b) Locust placed in manifold.

2.2.4 EAG normalization and analysis

To compare EAG response magnitudes, first data was sub-sampled to 100 Hz. Next, DC offsets were removed by subtracting mean pre-stimulus for each intensity-humidity pair. For comparisons across locusts, all data was normalized to the maximum magnitude of hexanol 1%

in 0% RH. Means and S.E.M. were plotted to compare the effect of intensity and RH on EAG response and statistical significance for changes with intensity were calculated using a two-tailed t-test with a Bonferroni corrected α of 0.05. Significance for changes in magnitude under different humidity backgrounds were calculated using a one-tailed t-test with a Bonferroni corrected α of 0.05.

To compare EAG baselines, data was sub-sampled to 100 Hz. Next, DC offsets were aligned by subtracting mean pre-stimulus for each stimulus in 0% RH. Significance for changes in baseline under different humidity backgrounds were calculated using a one-tailed t-test with a Bonferroni corrected α of 0.05.

2.2.5 Projection neuron recordings

Extra-cellular recordings were performed using previously published methods [11-13]. In brief, a young adult locusts of either sex was secured to a platform and a cup of wax was built around its head. The cup was filled with locust Ringer's solution (locust saline [16]) and a small section of the locust skull and underlying support tissue was removed. Approximately 0.3-0.4 mg of protease (P5147, Sigma-Aldrich. St. Louis, MO) was applied to sheath directly over the antennal lobes for 5-10 seconds then thoroughly rinsed and removed with locust saline. A small window was then excised in the sheath directly above each antennal lobe. For recordings, a drip line of locust saline flowing at approximately 50 ml/min was added.

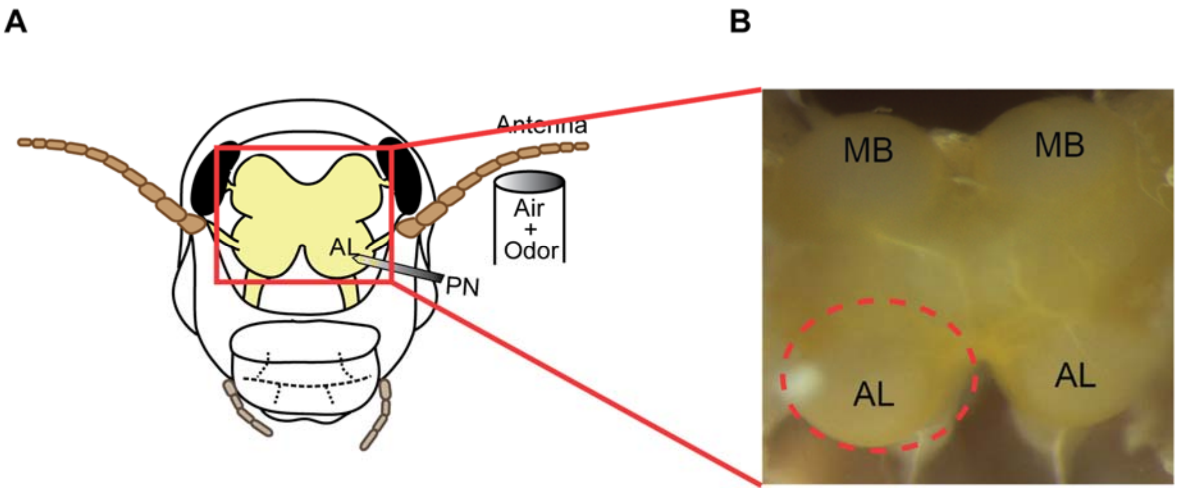


Figure 2.14. Projection neuron recording in locust antennal lobe. A) Schematic image of PN recording preparation. Window opened in skull allows for access to antennal lobe and insertion of electrode. B) Image of locust brain. Antennal lobs are labeled as AL.

A ground wire was placed in the saline well and a 4x4 16 channel probe (A2x2-tet-3mm-150-121, NeuroNexus) was inserted through the window in the sheath so that the electrode pads were just touching the surface of the antennal lobe (Figure 2.14). The preparation was allowed to rest for at least 15 minutes before recordings were started to allow the electrode to stabilize in the tissue.

Prior to insertion, the electrode was gold plated to obtain impedance values between 200-300 kΩ. The extracellular signals were amplified 10,000X, bandpass filtered between 0.3-6 kHz, and sampled at 15 kHz.

2.2.6 Spike sorting

Individual projection neurons were identified from our extracellular using a published spike sorting technique previously used in our work [11-13, 17]. Briefly, we used the following criteria in order to sort individual spikes: cluster separation was greater than 5 times the noise standard deviation, the number of spikes within 20 ms had to be less than 6.5% of the total spikes, and the spike waveform variance was less than 6.5 times the noise standard deviation.

Two separate datasets were collected. One for 4 odorants at a 1% (v/v) concentration yielding 82 PNs from 19 locusts. The other for hexanol at 3 different concentrations (0.1%, 1%, and 10% v/v) yielding 94 neurons from 17 locusts.

2.2.7 Peri-stimulus time histograms

Actual potentials were summed in 50 ms non-overlapping bins and smoothed by a 5 or 10 point average zero-phase digital filter. Averages across trials were computed for single unit PSTHs and averages across all trials are neurons were calculated for population-level PSTHs.

2.2.8 Responsive PN selection

PNs were categorized as selective based on two criteria: 1) Activity criteria: In at least one time bin during the stimulus, activity had to exceed 3 S.D. of the average, pre-stimulus activity. 2) Consistency criteria: The activity criteria had to be met in at least 3 of the 5 trials.

2.2.9 Neural correlations

Correlations for the ensemble PN activity were calculated as follows. Using a custom MATLAB (vR2016B, MathWorks, Natick, MA) script, the mean firing rate during the 4 seconds of stimulus was calculated for all PNs and assembled in a vector. Correlations of the vectors were calculated for two cases: 1) across odorants and humidity levels and 2) across intensities and humidity levels.

2.2.10 Dimensionality reduction analysis

We used a linear principal component analysis technique for the purpose of visualizing high-dimensional neural response trajectories. For this analysis, we binned the PN responses in 50ms non-overlapping time bins and averaged the responses in a given time bin across trials. PN responses were thus arranged as time series data of n dimensions (where n is number of neurons recorded) and m steps (the number of 50 ms time bins). Each high dimensional PN response vector in each time bin was projected onto the three leading eigenvectors of the response covariance matrix. For the analysis shown in Figure 4.3 D, we obtain covariance matrix from response of four odorants and all humidity conditions. The covariance matrices used in Figure 4.5, however, were obtained separately for each odorant. Finally, the low-dimensional points were connected in a temporal order to generate low dimensional response trajectories. The response trajectories were smoothed using a five-point moving average filter.

2.2.11 High dimension neural classification

High dimensional classification was done as in previous studies [12, 13]. Neural responses were binned in 50ms non-overlapping time bins to generate ensemble neural response profiles over time. Three, trial and time averaged, reference templates were generated for classification, one for each hexanol concentration for activity during the first 2s following onset of odor delivery. All reference templates were generated from activity in the 0% RH condition.

Classification was performed using an angular distance metric. Response vectors with the smallest distance to a reference template were classified as belonging to the corresponding class. For classification of 0% RH trials, a leave-one-trial-out validation was used. To ensure only meaningful responses were classified, a tolerance threshold of 63° was used requiring all classified response vectors to be within a certain distance of the reference templates.

2.3 Locust behavior

2.3.1 Locust preparation

Locust behavior experiments were performed following previously published methods [12, 13, 18]. In brief, adult locusts of either sex were starved for 24 hours and then immobilized in plastic syringes with an opening cut to ensure the locust antennae, head, and mouthparts were visible and accessible. The eyes were closed with a black tape and the distal segments of the maxillary palps (mouthparts) were painted green for image tracking purposes with a zero volatile organic compound paint (Valspar Ultra). The training sessions began approximately an hour after the palps were painted to allow proper drying and to allow the locusts to adjust to being immobilized and have paint on their palps.

2.3.2 Locust training and testing

1% Hexanol was used as the conditioned stimulus (CS) in training and organic wheat grass was the unconditioned stimulus (US). Training was performed in either a 0% or 100% RH carrier air stream. Odorant delivery and humidity control were achieved as described above. In brief, locusts were exposed to carrier air at 750 sccm at all times. Odor pulses were presented in addition to the air at 100 sccm. To clear delivered odorants, a vacuum line was placed behind the locusts. A video camera (Microsoft Webcam) was used to record the locusts' palp opening response (POR) at 30 frames per second. Odor delivery and video data acquisition were automated using custom written LabVIEW 2009 (National Instruments, Austin, TX) programs for precise testing conditions.

Locusts were trained over 6 training trials, spaced 10 minutes apart, in which the CS was presented for 10 seconds and the US was presented 5 seconds after the start of the CS for approximately 10 seconds. Only locusts which ate wheat grass in at least four out of the six training trials and performed a satisfactory palp-opening response (POR) in at least three of the six trials were retained for the testing phase (67% of the locusts in 100% RH (20/30) training and 73% of the locusts in 0% RH (22/30) training, note that 2 trained in the dry condition were later removed due to poor palp-tracking).

In the testing phase (results reported in Figure 8), locust PORs were collected for a total of 6 trials. The 6 trials were a pseudo randomized combination of 0.1%, 1%, and 10% hexanol being presented for 4 seconds in the presence of 0% or 100% RH air. The inter-trial delay was set to 25 minutes. Locusts were kept on a 12 h day – 12 h night cycle (7 am – 7 pm day). All behavioral experiments were performed between 9 am – 3 pm.

2.3.3 Palp-tracking algorithm

Maxillary palp movement videos acquired from LabVIEW were analyzed using a custom written Matlab script. The goal of the processing was to provide contrast to enable accurate tracking of the painted palps. Each trial was saved as a separate 30 second video that recorded both pre- and post-stimulus palp movements. The videos were cropped to focus on the region of the palps to minimize background noise and then converted into a time series of RGB color frames. It was ensured that the videos still contained 30 frames per second and that no data had been lost.

For each frame of a video, the grayscale image was subtracted from the green channel. A 2-D averaging filter (10 pixels by 10 pixels) was then applied to remove background noise and emphasize the green painted palps. Since no two locusts were identical, a number of parameters were manipulated to accommodate for differences in palp size, position, movement intensity, and visibility of the painted region as the palps moved. Details of these procedures have been previously published [13].

Despite these manipulations, if the program lost track of a palp for a frame due to antennal movements blocking the view or low pixel intensity, it would estimate the position of the palp using its position in the previous frame. The position of the centroid of each painted palp was tracked in every frame and the distance between the centroids was used as an estimate of the distance between the palps throughout the trial. The results of each video analysis were manually inspected and the program was re-run until the results reflected the movements accurately. The distance curve obtained was then smoothed using an 8 point zero-phase digital filter. 1 of the 22 locusts that were trained in 0% RH and subsequently tested was excluded from analyses due to rapid antenna movements that hindered accurate palp tracking. Another was removed due to poor palp tracking, due to insufficient paint being applied. All 20 locusts trained in 100% RH were included.

2.4 Identifying malaria biomarkers

2.4.1 Breath collection sampling methodology

Alveolar air samples were collected form patients as detailed in other studies [19]. At least 1 L of breath was collected in 3 L sample bags (SamplePro Flexfilm sample bags, SKC Inc.). To transport samples back to St. Louis for analysis, exactly 1 L of the collected breath was pumped through an inert stainless steel sorbent tube (Tenax 60/80/ Carbograph 1 60/80/ Carboxen 1003 40/60, Camsco). The use of a set flow pump (ACTI-VOC, Markes International) calibrated to 100 mL/min ensured the correct volume from the collection bag was sampled.

2.4.2 Data extraction and pre-processing

Upon return, samples were desorbed from the collection tubes using a TurboMatrix 650 ATD (Perkin Elmer) and analyzed with a Leco Pegasus 4D GCxGC-TOFMS system. Details of the analysis are detailed in our manuscript [20]. A brief overview of the method is as follows. Prior to desorption a standards mixture of known quantities of flourobenzene, tolune-D8, bromofluorobenzene, and 1,2-dichlorobenzene-D. Samples went through a 2-phase desorption process before being injected into the GC/MS. The GC column was a DB-5 column of 30m length, 0.25 mm ID, and 0.25 μm film thickness. GC spectra were analyzed and base ion peaks were normalized to the base ion peak area of the 1,2-dichlorobenzene-D4 standard. An aligned, standardized compound list generated by Mass Profiler Professional was exported and internal standards and silicone/siloxane contaminants were manually removed. A representative GC spectra at different stages of processing is shown in Figure 2.15. Note the change in spike abundance and density as the spectra is normalized and contaminants are removed (Fig. 2.15 A-B). Peaks present in only one patient were removed and the spectra were put into a log2 scale for analysis.

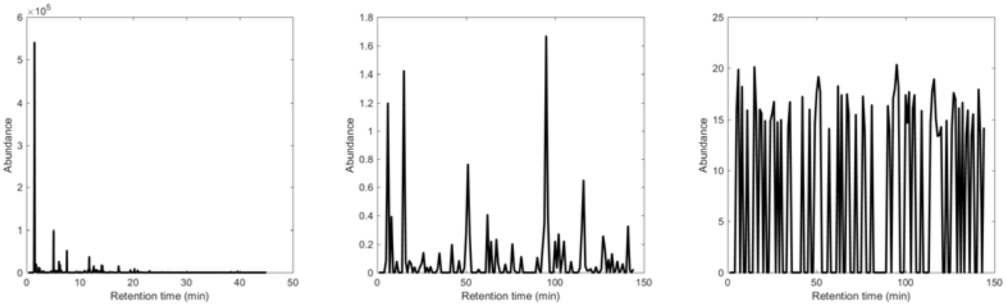


Figure 2.15. Representative GC/MS spectra collected from pediatric patients in Malawi. A) Unprocessed GC/MS spectra. B) Spectra normalized to 1,2-dichlorobenzene-D4 internal standard with siloxane contaminants removed as well as compounds present in only 1 patient. C) Logarithmic compression of spectra from b used to emphasize differences.

2.4.3 Cumulative abundance classifier

The aligned, standardized compound list was used for further analysis to determine malaria biomarkers. Only VOCs that were present in at least 20 participants were used in this analysis, leaving 42 potential biomarkers of 144 identified after contaminant removal. Class labels were assigned to each subject based on their diagnosis, as follows: 1 for malaria negative and 2 for malaria positive.

First VOCs were sorted based on their correlation with class labels. Higher correlation magnitudes indicated greater suitability as a biomarker for malaria. To classify a subject, the abundances of the six most correlated VOCs were summed in order of correlation magnitude. For negatively correlated VOCs, abundances were subtracted from the overall cumulative value. A nearest mean classification algorithm (binary classification) with leave-one-breath-sample-out cross validation scheme was followed. In this approach, mean cumulative VOC abundances for both classes were computed, excluding the breath sample from the subject to be classified. The left out breath sample was classified based on the class mean to which it was closest. This was repeated systematically leaving out each sample in the dataset. The predicted label [malaria (+) or malaria (-)], was compared with the actual values in order to quantify the performance as shown in Fig. 5.7C-D. A detailed characterization of how the performance varies with the number of VOCs selected is provided in Fig. 5.4.

The classification performance as a function of number of VOCs included was used to determine the optimal number of VOCs needed for identification of subject disease state. Based on this metric, six VOCs were selected as promising components for a diagnostic breathprint.

2.5 References

1. Katta, N., et al., *The I/O transform of a chemical sensor.* Sensors and Actuators B: Chemical, 2016. **232**: p. 357-368.

2. Benkstein, K., et al., *Integration of nanostructured materials with MEMS microhotplate platforms to enhance chemical sensor performance.* Journal of Nanoparticle Research, 2006. **8**(6): p. 809-822.

3. Martinez, C.J., et al., *Porous Tin Oxide Nanostructured Microspheres for Sensor Applications.* Langmuir, 2005. **21**(17): p. 7937-7944.

4. Semancik, S. and R. Cavicchi, *Kinetically Controlled Chemical Sensing Using Micromachined Structures.* Accounts of Chemical Research, 1998. **31**(5): p. 279-287.

5. Semancik, S., et al., *Microhotplate platforms for chemical sensor research.* Sensors and Actuators B: Chemical, 2001. **77**(1–2): p. 579-591.

6. Shirke, A.G., et al., *Femtomolar isothermal desorption using microhotplate sensors.* Journal of Vacuum Science & Technology A, 2007. **25**(3): p. 514-526.

7. Cavicchi, R.E., et al., *Featured Article: Use of Microhotplates in the Controlled Growth and Characterization of Metal Oxides for Chemical Sensing.* Journal of Electroceramics, 2002. **9**(3): p. 155-164.

8. Rogers, P.H. and S. Semancik, *Development of optimization procedures for application-specific chemical sensing.* Sensors and Actuators B: Chemical, 2012. **163**(1): p. 8-19.

9. Raman, B., et al., *Designing and optimizing microsensor arrays for recognizing chemical hazards in complex environments.* Sensors and Actuators B: Chemical, 2009. **137**(2): p. 617-629.

10. Wang, W.-N., Y. Jiang, and P. Biswas, *Evaporation-Induced Crumpling of Graphene Oxide Nanosheets in Aerosolized Droplets: Confinement Force Relationship.* The Journal of Physical Chemistry Letters, 2012. **3**(21): p. 3228-3233.

11. Saha, D., et al., *Multi-unit Recording Methods to Characterize Neural Activity in the Locust (Schistocerca Americana) Olfactory Circuits.* 2013(71): p. e50139.

12. Saha, D., et al., *A spatiotemporal coding mechanism for background-invariant odor recognition.* Nat Neurosci, 2013. **16**(12): p. 1830-1839.

13. Saha, D., et al., *Behavioural correlates of combinatorial versus temporal features of odour codes.* Nat Commun, 2015. **6**.

14. Saha, D. and B. Raman, *Relating early olfactory processing with behavior: a perspective.* Current Opinion in Insect Science, 2015. **12**: p. 54-63.

15. Boeckh, J., K.E. Kaissling, and D. Schneider, *Insect Olfactory Receptors.* Cold Spring Harbor Symposia on Quantitative Biology, 1965. **30**: p. 263-280.

16. Laurent, G. and M. Naraghi, *Odorant-induced oscillations in the mushroom bodies of the locust.* The Journal of Neuroscience, 1994. **14**(5): p. 2993-3004.

17. Pouzat, C., O. Mazor, and G. Laurent, *Using noise signature to optimize spike-sorting and to assess neuronal classification quality.* Journal of neuroscience methods, 2002. **122**(1): p. 43-57.

18. Simões, P., S.R. Ott, and J.E. Niven, *Associative olfactory learning in the desert locust, Schistocerca gregaria.* The Journal of Experimental Biology, 2011. **214**(15): p. 2495-2503.

19. Berna, A.Z., et al., *Analysis of Breath Specimens for Biomarkers of Plasmodium falciparum Infection.* The Journal of Infectious Diseases, 2015. **212**(7): p. 1120-1128.

20. Schaber, C., et al., *Breathpriting Reveals Malaria-Associated Biomarkers and Mosquito Attractants.* Manuscript in preparation, 2017.

Chapter 3: The I/O transform of a chemical sensor

3.1 Introduction

Chemical sensing involves detection of a target analyte of interest by measuring a change in a signal that is generated by either the analyte's contact or remote interaction with the sensor. The minimum signal that evokes a measurable output above noise levels is the sensor's detection limit, and the minimum change in the signal levels that can be distinguished is its resolution or sensitivity. In this context, a sensor's operation is akin to a mathematical input/output (I/O) transform [1, 2]. Ideally, the sensing operation will be meaningful if this transform is different for different inputs. Assuming that this hypothesis is true, the proposed formalism faces several practical challenges:

i. how does one estimate the I/O transform of a sensor?
ii. how robust is the I/O transform to changes of parameters relevant to the sensing task such as identity and intensity of the target signal?
iii. how invariant is the I/O transform to changes in irrelevant parameters such as sensor age[3-8]?, and finally
iv. how reproducible are the I/O transforms for different sensors of equal manufacture?

We examined these issues in the context of chemical sensing with an array of chemiresistive microsensors.

Current approaches for chemical sensing have been inspired by the biological principle of using an array of cross-selective chemical sensors to create unique multivariate fingerprints for different odorants. A number of sensing technologies have been proposed for sensitive detection and selective recognition of chemicals [9-20]. Irrespective of the transduction mechanism used, all chemical-sensing approaches must employ a readout mechanism that would allow extraction of meaningful features from the sensor's response. Two strategies seem popular depending on the segment of the response that is assumed to contain most discriminatory information: transient or steady-state responses.

For steady-state response analysis, the ultimate change in the sensor signal after exposure to an analyte is typically the measurement used for further processing [8, 21-24]. In the case of response transients, time-domain features such as sensor‘s response time constants [25-29] or frequency-domain features such as spectral content [30-32] have been popularly used. In general, it is widely accepted in both artificial and biological chemical sensing systems that the sensor response during transient periods tends to carry richer analyte specific information and therefore can provide better recognition performances[28, 33-44].

Irrespective of the signal readout from the sensor (steady-state vs. transient or time-domain vs. frequency domain), another issue faced by almost all chemical sensors that severely limits their long-term use is the issue of sensor drift, or deviation of the response, over time[3, 7, 26, 45-49]. Drift in artificial chemical sensors has been suggested to primarily be an effect of aging or poisoning of the sensing film [5]. Drift has a profound effect on absolute transducer measurements, making these measures unreliable for long term analyte identification. Therefore, the long-standing need for the development of portable, accurate, and precise chemical sensors, which remain viable for extended periods of operation, still remains unmet.

In this work, we propose a simple but elegant approach to characterize the sensor's operation as an I/O transform. The proposed approach exploits the richness of a sensor's response to a temporally structured pulsatile sampling of analytes similar to those used in biological olfaction [37, 38, 41, 50]. It is worth noting that the sensor responses following stimulus onsets, offsets and during their steady-state are all used to characterize a sensor's response to an analyte. We reveal that this approach, to a certain degree, is robust with respect to changes in analyte intensity and to changes due to sensor drift. More importantly, we also show that this approach facilitates generalization between sensors of equal manufacture and therefore provides a way to allow seamless replacement of sensors in a chemical sensing system. The latter problem is fundamental for successful deployment of sensors in commercial applications.

3.2 Results

3.2.1 Responses of a chemical sensor to pulsatile stimuli

We began by examining the response of a metal-oxide chemiresistor to a few analytes that varied in their functional group, carbon chain length, and vapor pressure. Unlike most chemical sensing studies, we delivered stimuli in a pulsatile fashion to exaggerate the response transients. The pulses delivered varied in duration, (1 to 3) min, and inter-pulse interval, (1 to 4) min. The mean response of the sensor (±standard deviation (S.D.)) at one particular operating temperature (435 °C; see Fig. 2.5) is shown in Figure 3.1. Note that sensor responses to the exposed gas were measured at twenty eight operating temperatures (Fig. 2.5).

In general, irrespective of the operating temperatures, we found that all gases used increased sensor's conductance (i.e. all were reducing gases) with the absolute magnitude of the response being greater for ethanol and acetone (blue and green) compared to the others. We found that the pulsed stimulus sequence emphasized differences in the transient responses generated by different analytes. During a relatively lengthy pulse, there was a discernable difference in time to peak response between different analytes (Fig. 3.1, inset 1). Similarly, when the inter-pulse interval was reduced, responses to the non-leading pulses decreased substantially for all analytes (Fig. 3.1, inset 2). Note that the magnitude of response reduction was analyte specific. Hence, we hypothesized that the pulsed mode of stimulus exposures could enhance discriminability between analytes. Furthermore, since some of the response dynamics were governed by the differences in stimulus dynamics, we expected these transient portions to remain invariant to sensor drift, consistent with previously shown results [51].

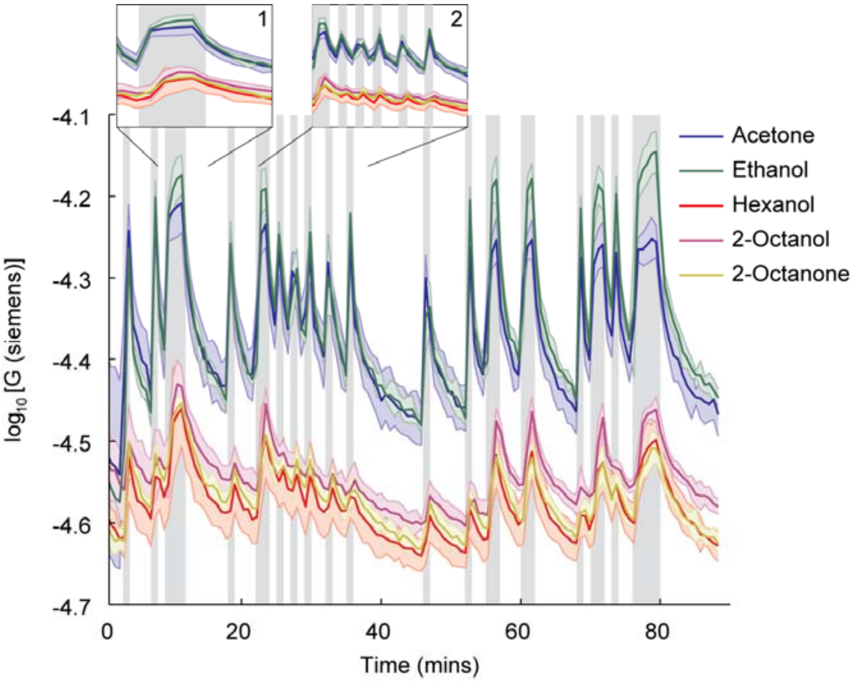


Figure 3.1. Responses of a metal-oxide sensor (SnO_2) to five different analytes presented in pulsatile fashion. Each trace represents the mean sensor response to an analyte with the sensor operating at 435 °C. Color bands represent standard deviations. Gray bars indicate periods when the sensor was exposed to the analyte. First inset reveals response onset differences during a prolonged single pulse. Second inset reveals diminishing of responses when probed with short, back-to-back stimulus pulses.

3.2.2 Chemical sensing as an I/O transform

For any given analyte, and at any particular point in time, the state of the valve delivering the stimulus (open or closed) and the sensor's response were both known. We viewed the sensor's operation as a transform that when applied to the stimulus state produced a corresponding response. If the sensor was extremely rapid and there was no hysteresis, then the I/O transform could be regarded to be instantaneous. However, this was not true for our sensor and for most chemical sensors studied in general. Therefore, we assumed that the sensor response at any given time is not instantaneous but a weighted linear function of the recent stimulus history. Note that this I/O operation or transformation must still be unique for each analyte in order for the sensor to generate unique fingerprints.

To estimate the linear transform from the training data, we recorded the stimulus history, as determined by the valve state, in a given time window ($\vec{x}$) and the sensor response at the end of this period (y; a scalar value). For example, the following valve-states for last eight measurements [closed$_{t-7}$, open$_{t-6}$, open$_{t-5}$, open$_{t-4}$, closed$_{t-3}$, closed$_{t-2}$, closed$_{t-1}$, closed$_{t}$], would result in the following stimulus history vector [0, 1, 1, 1, 0, 0, 0, 0]. For each stimulus and each trial, or run, we systematically moved the stimulus history window in time to generate a matrix of stimulus histories (X; 8 columns but multiple rows) and a vector of sensor responses measured at the end of each stimulus history (Y; column vector with the same number of components as the rows of X). The sensor response model to a given analyte could now be viewed as the optimal linear transform that reconstructs Y given X (schematically shown in Fig. 2.6). Since this results in an over-determined system of equations (more rows than columns), we calculated the I/O transform to minimize the mean-squared error of reconstruction (i.e. pseudo-inverse solution).

The optimal linear transform ($\vec{k}$; vector of the same dimensions as $\vec{x}$), and the comparison between the actual vs. reconstructed sensor responses for the different analytes used are shown in Fig. 3.2A and 3.2B. Note that each component of the I/O transform is a weight for the stimulus/valve state at a particular point in time/history. The overall sensor's response that is reconstructed in this fashion is merely a weighted sum of valve states in the recent past. As can be observed, the actual sensor response (black traces) and the predicted sensor response (red traces) are well-matched for all analytes (see Fig. 3.3).

The I/O transform of each analyte for each of the 28 operating temperatures used in the study is shown in Fig. 3.2C.

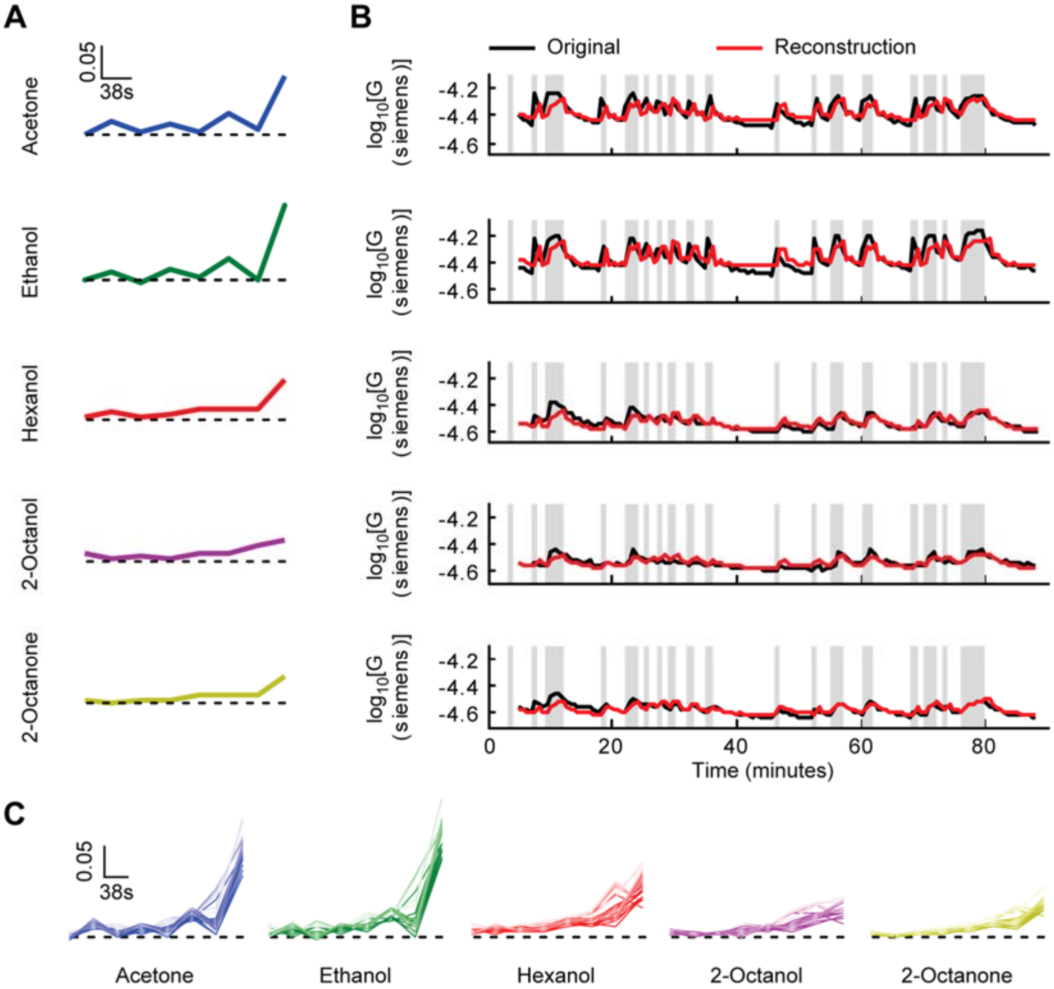


Figure 3.2. I/O transforms of a metal-oxide chemiresistor. A) I/O transforms ($\vec{k}$) generated using sensor's response at 435 °C for each of the five analytes examined in the study. C) Comparison of the sensor's actual response at 435 °C (black) and reconstructions (red) obtained for all five analytes. D) I/O transformations for each analyte at all twenty-eight temperatures used (see Figure 2.5). Higher intensity of color indicates higher temperature.

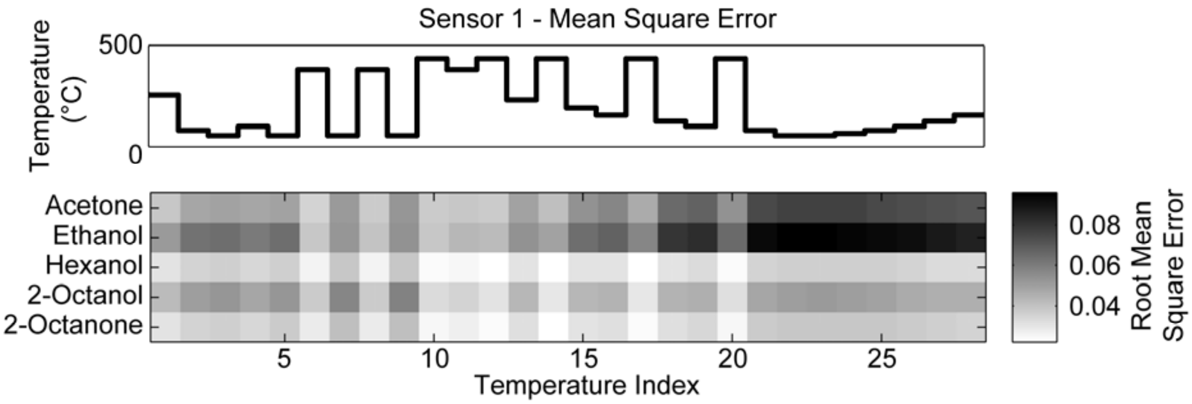


Figure 3.3. Root mean squared error for prediction of analyte response at each temperature in the cycle. Note that error is typically lower at higher temperature bands.

3.2.3 Recognition of chemicals based on Sensor's I/O transform

As can be noted from Fig. 3.2, the I/O transform was unique for each analyte examined. Furthermore, as shown in Fig. 3.4a, the I/O transform was reliable across multiple training runs for each analyte used. These results suggest that the estimated I/O transform of a sensor may be used as a fingerprint to identify each analyte. To confirm this hypothesis, we first visualized the eight-dimensional I/O transform using principal component analysis (Figure 3.4b). Note that each complete training run or trial, resulted in a single estimate of the I/O transform, and therefore is represented as a single point/symbol after PCA dimensionality reduction. Different runs corresponding to a particular analyte generally clustered together and were discriminable from clusters representing other analytes. Therefore, this result suggests that the sensor's I/O transform can indeed be used for analyte recognition.

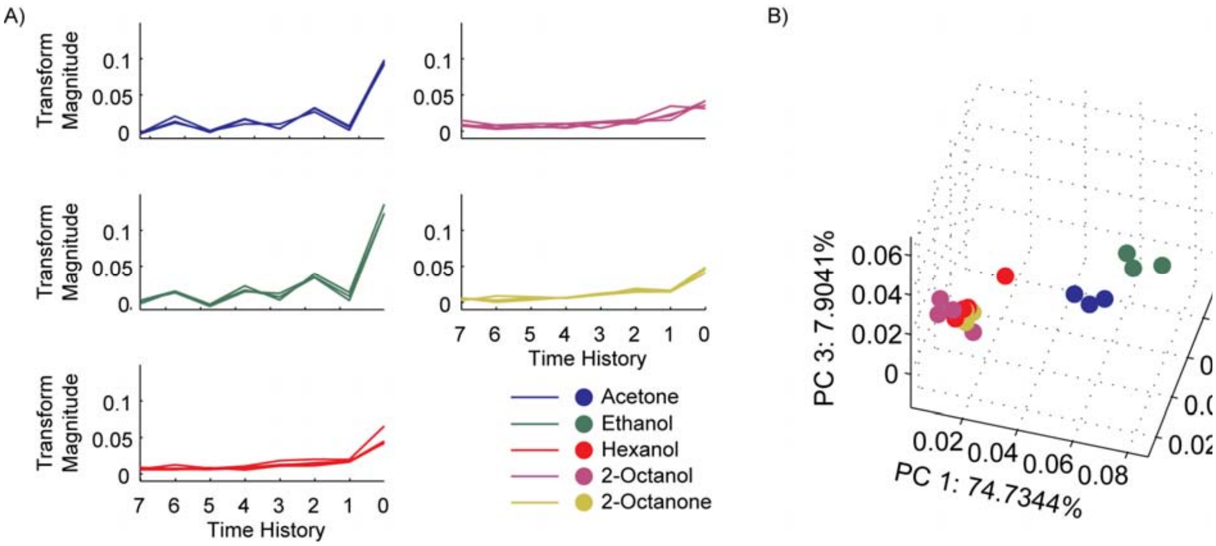


Figure 3.4. I/O transforms for different analytes are consistent across different training trials. A) I/O transforms obtained using sensor response at 435 °C are shown for each analyte and for each training run. B) Visualization of the I/O transforms ($\vec{k}$) of different analytes and for different runs using principle component analysis is shown.

3.2.4 Sensor's I/O transforms are drift tolerant

As mentioned before, drift can be a major issue with chemical sensors, as it can significantly reduce the viability of sensors operating over an extended period of time. To test the stability of the sensor's I/O transform to each analyte, we aged the sensor for a period of two months (see 2.1.5: second dataset). Subsequent exposures of the same five analytes generated raw sensor responses that were substantially drifted. We note that sensor baseline response decreased and the sensor response magnitude varied as a result of this drift (Figure 3.5).

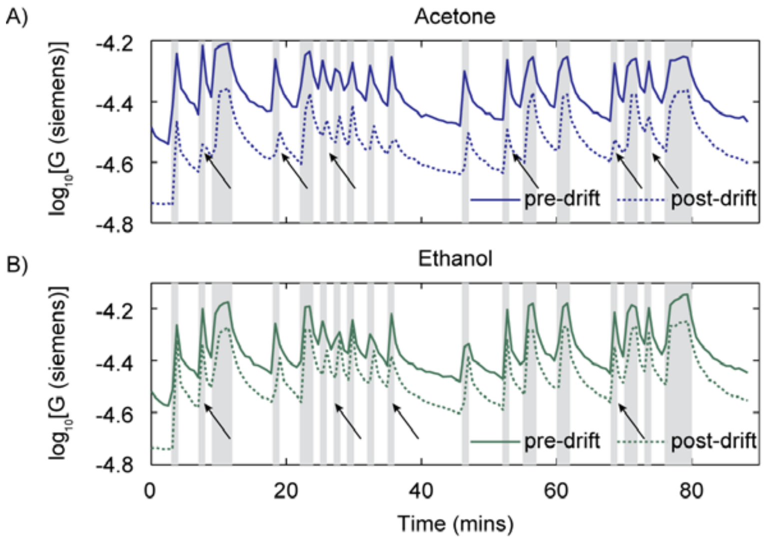


Figure 3.5. Sensor 1 response to each analyte before and after (2 to 3) months of aging. Each trace represents the mean sensor response to an analyte with the sensor operating at 435 °C (temperature index 20). Responses to acetone and ethanol vapors (a and b, respectively) are shown. Note the decrease in conductance caused by sensor aging. Although the traces look qualitatively similar, the black arrows indicate regions where relative response magnitude differs pre- and post- drift.

We made a qualitative comparison of sensor response profiles before (training dataset) and after (testing dataset) aging using a linear principle component analysis (Figure 3.6a). Note that sensor response from the training dataset is indicated using circles, whereas raw sensor responses during the validation phase are identified as squares. As can be noted, although groupings within datasets are well-defined, drift in sensor responses across datasets shifted the overall response profiles. Complementing this qualitative visualization analysis, we performed a quantitative nearest-neighbor classification. The overall results from this classification analysis are summarized in the confusion matrix shown in Figure 3.6b. Note that the main diagonal elements were low indicating misclassification. These results qualitatively and quantitatively confirm that the raw responses of the sensor before and after aging were inconsistent.

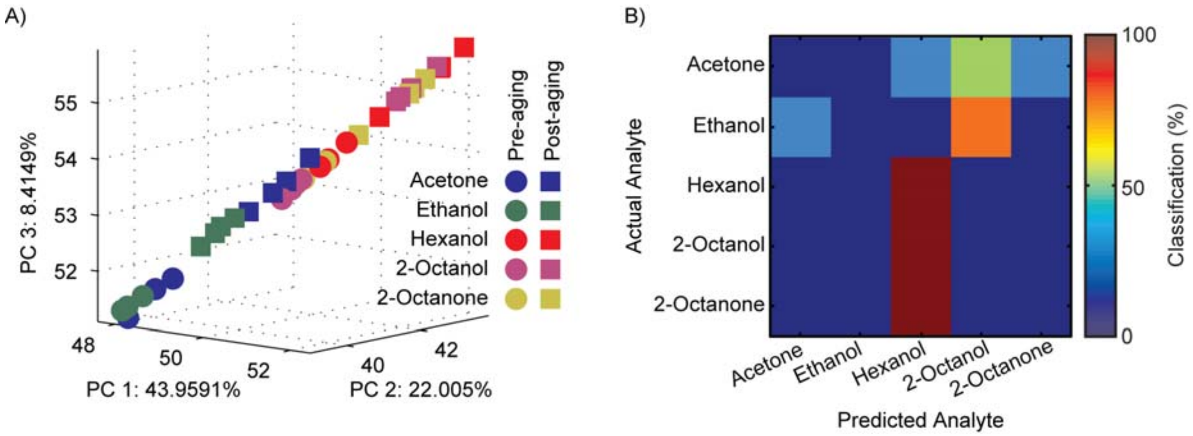


Figure 3.6. PCA clustering of sensor's data shows strong shift due to drift. A) Principle component analysis of sensor's responses to analytes at 435° C. The entire trace used to calculate the I/O transform was used to derive the three principle components. B) Classification of responses using same method as described in Figure 2.8.

Next, we compared the I/O transform for different analytes before and after sensor aging (Fig. 3.7). Despite the evident drift, the I/O transforms were relatively consistent (Fig. 3.7a). Further, to quantify the performance results we performed a classification analysis (Fig. 2.8). We reduced the overall classification problem into a series of two-class discriminations to determine similarity with ethanol vs. others (step1), acetone vs. remaining others (step2), 2-octanol vs. remaining others (step3), and finally hexanol vs. 2-octanone (step4). Note that this scheme allows progressive refinement at each step, focusing primarily on determining membership to the most distant response cluster (see methods). The discriminations are shown on the right in Fig. 2.8. The colored bands indicate regions, of the projection, where analytes would be successfully classified at that step and the gray bands indicate regions of the projection where analytes would be parsed through subsequent stages. The confusion matrix (Fig. 3.8A) shows that most of the analytes were recognized well above the chance level (17 % to 24 %), and significantly higher than a direct PCA approach (Fig. 3.8B).

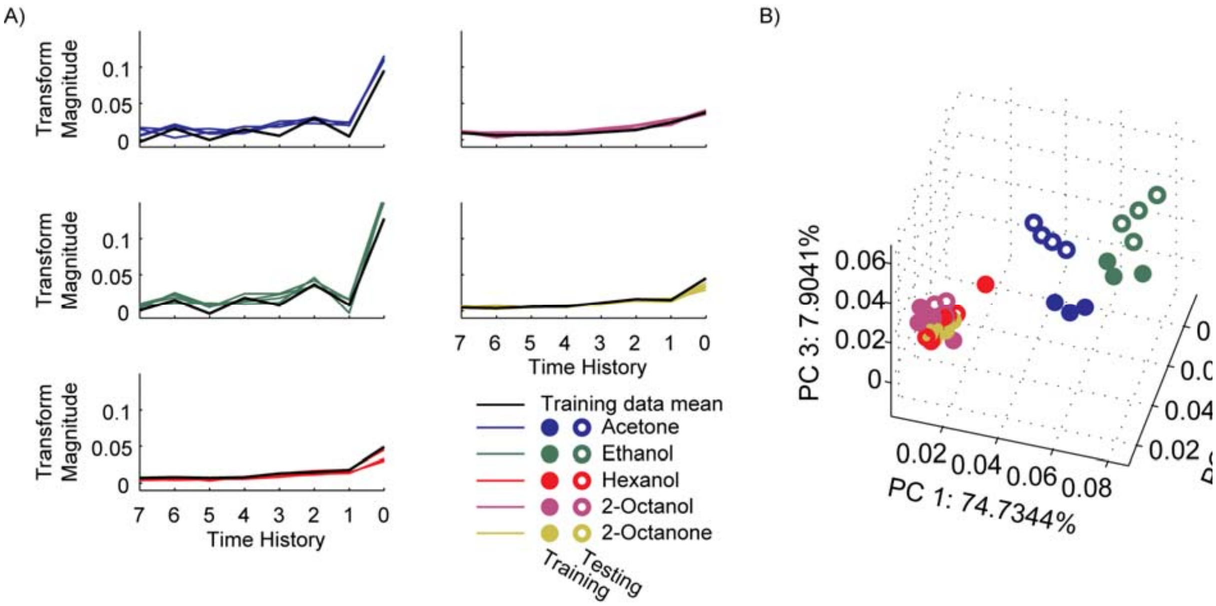


Figure 3.7. I/O transforms are robust with respect to sensor aging. A) All I/O transforms of sensor responses at 435 °C after aging are shown. Black line represents mean transforms obtained from pre-aged sensor (i.e. training data). Each colored line represents I/O transform obtained for an individual test phase trial. B) Principle component analysis of I/O transforms before (filled symbols) and after (open symbols) sensor aging. Only training data was used to calculate the principal component axes.

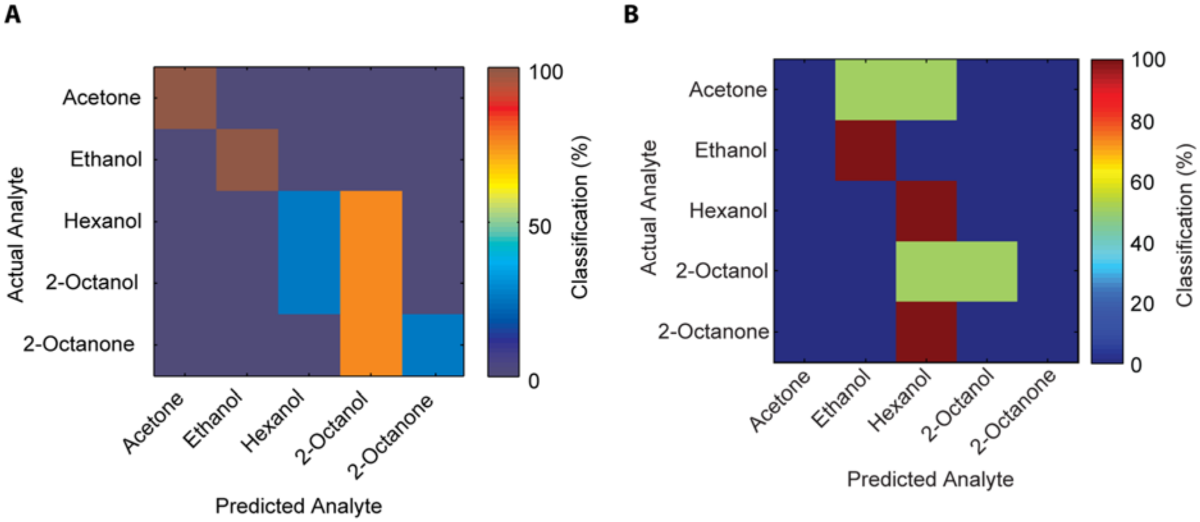


Figure 3.8. A hierarchical classification algorithm improves performance for multi-analyte recognition. A) Confusion matrix quantitatively summarizing the performance of the hierarchical classification approach. B) Confusion matrix quantifying classification performance using k-nearest-neighbors of only the dimensionality reduced space. Training data are from sensor 1 pre-aging. Testing data are from sensor 1 post-aging.

3.2.5 Concentration invariant recognition

Changes in concentration are also known to alter the sensor response magnitudes [52, 53]. Such response variations can lead to significant overlap in the responses generated by different analytes. Since the I/O transforms of the sensor to each analyte are predominantly focused on the response dynamics, we next examined how robust these were with respect to changes in stimulus intensity. We repeated the experiments with acetone, ethanol, hexanol, and 2-octanol but presented at different intensities.

We found that the I/O transforms were consistent even when the analyte concentrations were changed (Fig. 3.9a). In addition, we used a correlation based distant metric in our dimensionality reduction (Fig. 3.9b) to focus primarily in the shape of the I/O transform and not on its magnitude to further reduce sensitivity to variations that might arise due to stimulus intensity. As can be noted, the classification performance was well above the chance levels for all analytes tested in this fashion (Fig. 3.9c).

To further clarify these results, we scaled the response of a particular analyte using three different values (x1, x2, and x5). Such scaling provided responses that had similar transients and time constants, but with varying magnitude (Fig. 3.10). As expected, we found that the I/O transforms obtained for these three responses were identical in shape but differed substantially in their magnitude. Further, we note that a distance metric based on the correlation values would classify these I/O transforms to be identical thereby providing the analytical basis for invariance with respect to the changes in response magnitude.

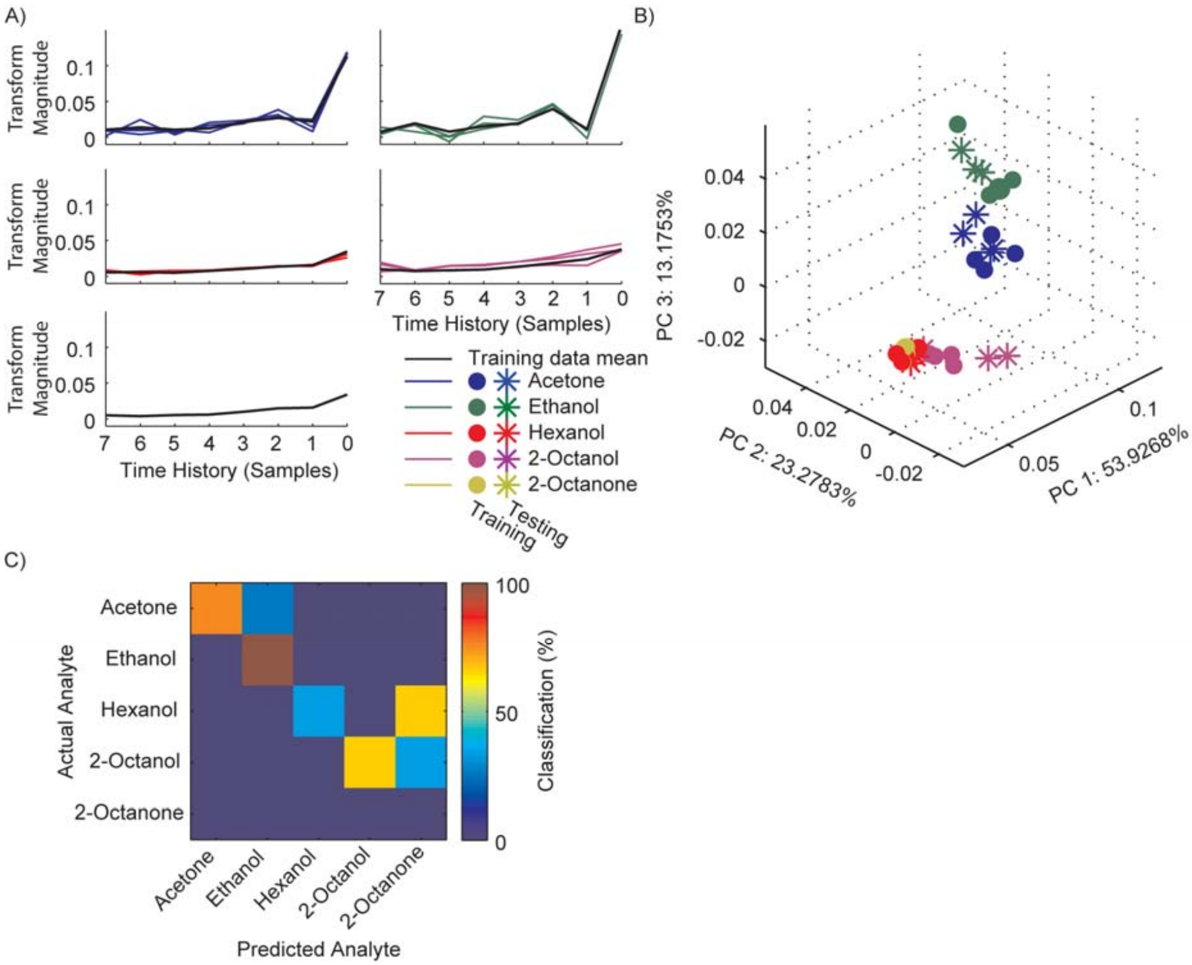


Figure 3.9. I/O transforms are robust with respect to analyte concentration. A) Each panel reveals I/O transforms estimated from a sensor's responses at 435 °C to analytes presented at varying concentrations. Black line represents mean transforms obtained from sensor response to a fixed concentration (i.e. training data). Each colored line represents an I/O transform obtained for an individual test phase trial. B) Visualization of I/O transforms obtained at varying concentrations of analytes are shown (filled symbols – training datapoints; asterisks – test datapoints). Only training data were used to calculate the principle component axes. C) Confusion matrix quantifying classification performance.

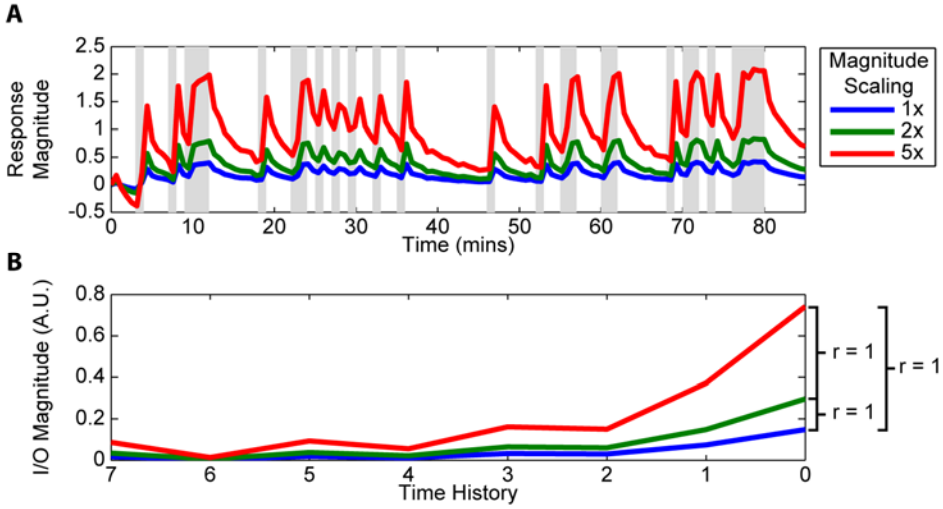


Figure 3.10. I/O transforms of scaled sensor responses are perfectly correlated. A) Response of the sensor to Ethanol vapors scaled using three different values (1x, 2x, and 5 x). B) Comparison of the I/O transformations generated for all three scaled versions of the ethanol responses are all perfectly correlated.

3.2.6 Sensor-invariant analyte recognition

Finally, we examined how robust the proposed signal extraction approach was across different equivalent sensors that were fabricated together. To examine this, we repeated the same analyses and compared the I/O transforms obtained between two sensors. We found that for each analyte, the I/O transforms were surprisingly consistent and training data from one sensor can allow identification of those analytes even when data from a different sensor was used for validation (Fig. 3.9). Note that these results are largely consistent with using training and testing datasets from the same sensor (refer Fig. 3.7). This further suggests that the response dynamics are dominated by the stimulus dynamics of the analyte themselves and may provide a robust approach for analyte recognition. Therefore, we expect this approach may also insulate the signal processing approaches to changes happening in the sensor array when damaged sensors are replaced with copies of the same type of sensor.

3.3 Discussion and conclusions

In this work, we have presented a method to characterize a chemical sensor's operation as a linear I/O transform. The input to the sensor is the recent stimulus history, which we defined here as a short time-series of the valve states ('ON' or 'OFF'). The sensor's response at the end of the stimulus history became the output to be predicted. In this formulation, we showed that for each analyte the sensor's operation became a unique input-output filter or a transform. Such an approach for developing a mathematical model of a dynamical system based on the inputs provided and the outputs measured is commonly recognized as a 'Systems Identification' approach [1, 2, 46, 54-56]. However, in the context of chemical sensing, what is not clear is how to use this system identification approach for recognition/differentiation of different analytes, how stable are these estimated I/O relationships, and how they vary over time, sensors etc. We have carefully explored these issues in this study.

In order for such a scheme to be feasible, it is first important to test whether the impulse response function of a chemical sensor can be reasonably estimated with a random inputs of certain finite length. If this assumption is reasonable, then the I/O transform (or alternately impulse response function) computed over one segment of random pulsatile binary inputs should allow prediction of the sensor's response to other random patterns of pulsatile binary inputs as well. We found that this is indeed feasible and the proposed estimation approach is robust and works well even when predicting the response of the sensor to other patterns of random binary input pulses (Fig. 3.11).

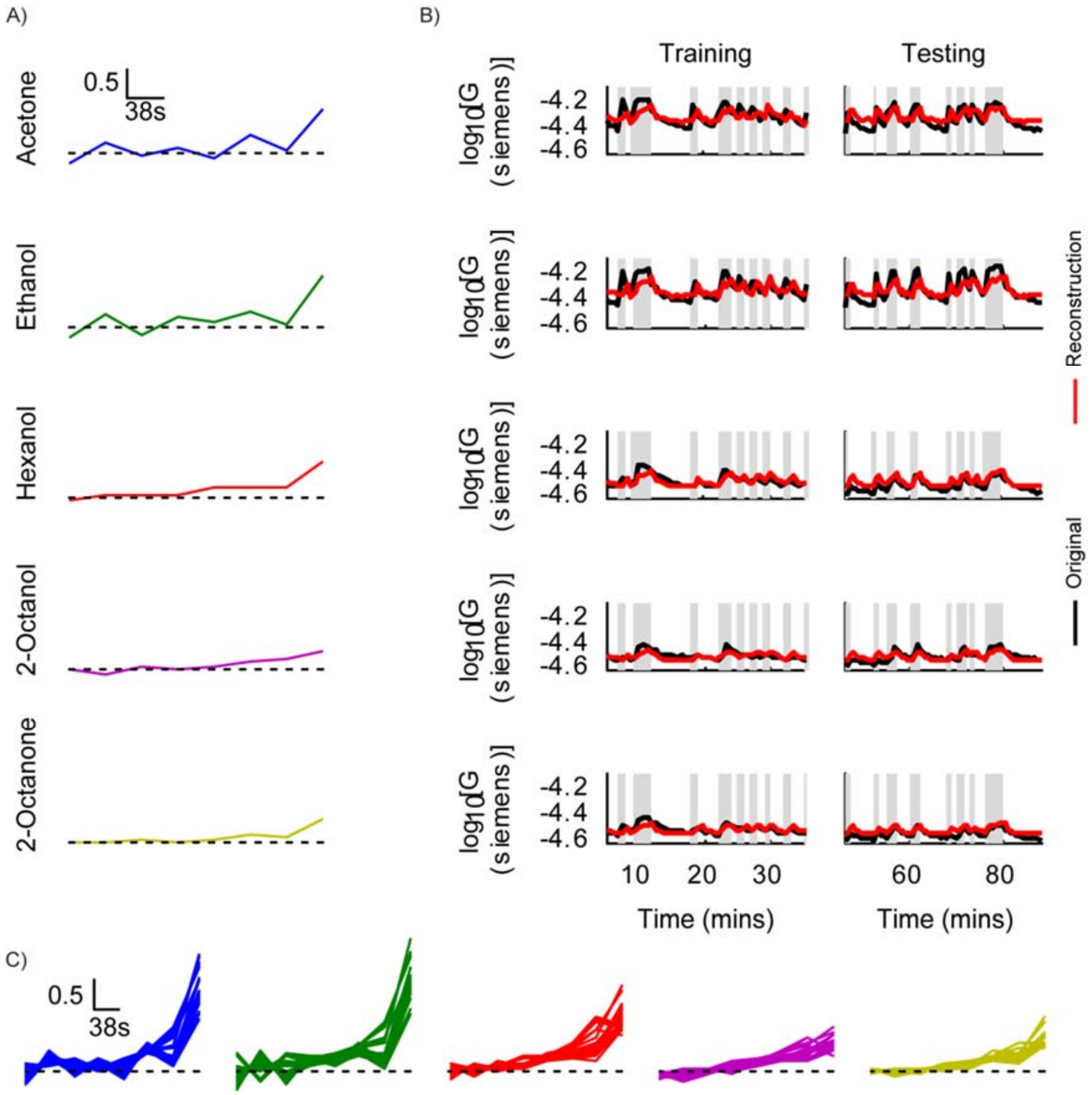


Figure 3.11**.** I/O transforms of a metal-oxide chemiresistor. A) I/O transforms ($\vec{k}$) generated using sensor's response at 435 °C for each of the five analytes are shown. Only half the entire stimulus sequence used in Figure 2.2 used to generate these transforms. B) Comparison between the sensor's actual response at 435 °C (black) and reconstructions (red) obtained for all five analytes for the first half (training) and second half (testing) of the stimulus sequence are provided. Note that the stimulus sequences during training and testing portions are different. C) I/O transformations for each analyte at all twenty-eight temperatures used are shown.

We note that the analyte discriminability was enhanced due to the employment of pulsatile stimulus delivery that enhanced information in the transient sensor responses. The response dynamics were largely driven by the stimuli themselves possibly due to differences in vapor pressure, the physisorption onto the sensor surfaces, etc. Therefore, we found that these analyte-specific transforms were robust, to a certain extent, to variations in analyte intensity and sensor age. The latter feature allowed reliable recognition of the analytes even when the sensor's baseline and the magnitude of stimulus-evoked responses changed over extended periods of operation (see Fig. 3.1, Fig. 3.5, Fig. 3.6). Hence, we believe that this approach may provide a drift-invariant analyte recognition scheme, a key requirement towards realizing non-invasive chemical sensing.

The only additional requirement imposed by the proposed technique is the need for active sampling approach as opposed to the typical passive methodology used for gas sensing. While other efforts that have examined the use of a pulsed stimulus delivery protocol for generating information rich datasets from the sensor, these works have either focused on the magnitude of the response [23] from a few short pulses or select features from the signal [57]. However, our approach markedly differs in that we examined not only the signal from stimulus exposure (onset transients and steady-states) but also took into account the transients generated following absence of a gas stimulus (i.e. offset transients). It will be worth noting that, although OFF-responses are informative, they are not considered by most approaches for discriminating analytes.

Such active sampling techniques are routinely used by biological systems to sample the chemical stimuli encountered in their environment ("antennal flicks" in invertebrates [58] or "sniffs" in vertebrates [59]). Active sampling is thought to extend greater control of the stimulus to the system, allowing it to manipulate and define the stimulus dynamics. Previous works have shown that stimulus dynamics can be dependent on the analyte or 'odorant' [37, 40] and may be exploited by subsequent processing centers in the brain [38, 41]. Several approaches to generate spatiotemporal sensor response profiles for analytes to enhance their discrimination have also been explored in artificial olfaction [39, 57, 60]. Our work complements these studies and focuses on development of schemes that take advantage of such rich data streams for the purpose of robust chemical identification.

Using such as scheme in a real-world scenario could be easily achieved by placing a small pump downstream of the sensor and drawing air carrying the encountered analyte over the sensor in a known binary pattern. In this scenario, the I/O transform can be estimated if the random binary stimuli used to control the pump/valve over a period of time and the output of the sensor during this active sampling period are both known. The estimated I/O transform of the currently encountered analyte can then be compared and pattern matched with those of the training analytes for the purpose of recognition.

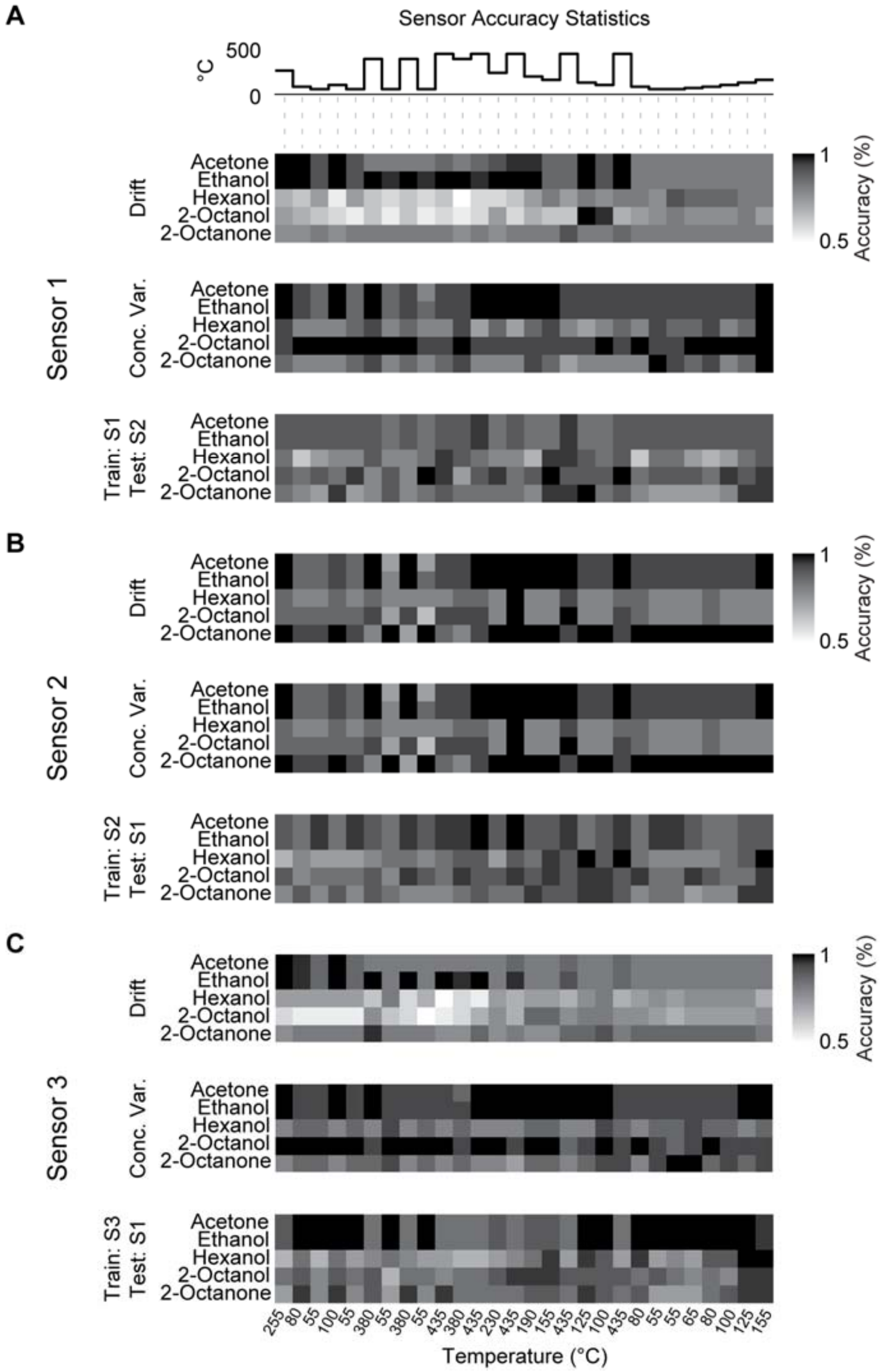


Figure 3.12. Classification performance is temperature and hysteresis dependent. A) Classification accuracy for sensor 1 is shown for each temperature in the cycle for the three conditions examined (from top to bottom): (2 to 3) months drift, for varying concentrations, and across different sensors. The temperature cycle used in the study is shown at the top of the plot. In general, for all three cases, the classification performance is better in the second half of the temperature cycle. B-C) Same as A but for sensors 2 and 3 respectively.

We note that the proposed scheme has not been optimized for rapid recognition of analytes. We envision changes that could be made to increase the sampling rate and therefore decrease the duration and spacing of pulses used to address this issue in future studies. Furthermore, in agreement with previous studies [26, 61], we found information content across temperatures was redundant. Hence, our temperature programs could be optimized by down-selecting to several high and low temperatures [56, 62-64]. We also found that hysteresis has a pervasive effect in metal oxide gas sensors. In general, our results show that I/O transforms of responses collected at temperatures above 250 °C, especially those that occur in the later segments of the temperature cycle, were more consistent across analytes and therefore resulted in higher classification accuracies (Fig. 3.12). In this study, we did not systematically attempt to take advantage of the hysteresis related effects we observed, which would provide another free parameter for the purpose of optimization.

Finally, we found that the sensors of equal manufacture generated strikingly similar transforms for a given analyte. This allowed the data obtained from one sensor to be used for recognizing the training analytes with a different sensor. The transferability of the training data is based upon the identification of sensor-independent features for pattern recognition. These can then be used to improve the flexibility of sensor arrays for a variety of application areas, primarily enabling greater longevity once deployed. Our demonstrated approach of pulsatile sampling and I/O transforms has the potential to be a simple, technology-independent technique for achieving this capability.

3.4 Author contributions

The work in this chapter has been published as: Katta, N., D. C. Meier, K. D. Benkstein, S. Semancik and B. Raman (2016). "The I/O transform of a chemical sensor." Sensors and Actuators B: Chemical **232**: 357-368.

Experiments and analysis were designed by NK and BR and were carried out by NK. Sensor array was provided by DM, KB, and SS. The manuscript was written by NK with revisions from BR. This work was funded by an Office of Naval Research grant (N00014-12-1-0089), NSF CAREER grant (1453022), and Children Discovery Institute’s Interdisciplinary Research Initiative grants to B.R.

3.5 References

1. Meade, M.L. and C.R. Dillon, *Signals and systems : models and behaviour*, in *Signals and systems : models and behaviour*. 1991, Chapman & Hall: London ; New York. p. 73-75.

2. Santiago, M., et al., *Different strategies for the identification of gas sensing systems.* Sensors and Actuators B: Chemical, 1996. **34**(1–3): p. 213-223.

3. Holmberg, M., et al., *Drift counteraction in odour recognition applications: lifelong calibration method.* Sensors and Actuators B: Chemical, 1997. **42**(3): p. 185-194.

4. Holmberg, M., et al., *Drift counteraction for an electronic nose.* Sensors and Actuators B: Chemical, 1996. **36**(1–3): p. 528-535.

5. Pearce, T.C., *Handbook of machine olfaction : electronic nose technology*. 2003, Weinheim Germany: Wiley-VCH. xxxii, 592 p.

6. Meier, D.C., B. Raman, and S. Semancik, *Detecting Chemical Hazards with Temperature-Programmed Microsensors: Overcoming Complex Analytical Problems*

with Multidimensional Databases.* Annual Review of Analytical Chemistry, 2009. **2**(1): p. 463-484.

7. Raman, B., et al., *Bioinspired Methodology for Artificial Olfaction.* Analytical Chemistry, 2008. **80**(22): p. 8364-8371.

8. Gutierrez-Osuna, R., *Pattern analysis for machine olfaction: a review.* Sensors Journal, IEEE, 2002. **2**(3): p. 189-202.

9. Grate, J.W., *Acoustic Wave Microsensor Arrays for Vapor Sensing.* Chemical Reviews, 2000. **100**(7): p. 2627-2648.

10. White, J., et al., *Solid-State, Dye-Labeled DNA Detects Volatile Compounds in the Vapor Phase.* PLoS Biol, 2008. **6**(1): p. e9.

11. Gopalakrishnan, D. and W.R. Dichtel, *Direct Detection of RDX Vapor Using a Conjugated Polymer Network.* Journal of the American Chemical Society, 2013. **135**(22): p. 8357-8362.

12. Goldsmith, B.R., et al., *Biomimetic Chemical Sensors Using Nanoelectronic Readout of Olfactory Receptor Proteins.* ACS Nano, 2011. **5**(7): p. 5408-5416.

13. Lu, C.-J., et al., *Portable Gas Chromatograph with Tunable Retention and Sensor Array Detection for Determination of Complex Vapor Mixtures.* Analytical Chemistry, 2003. **75**(6): p. 1400-1409.

14. Banan Sadeghian, R. and M. Saif Islam, *Ultralow-voltage field-ionization discharge on whiskered silicon nanowires for gas-sensing applications.* Nat Mater, 2011. **10**(2): p. 135-140.

15. Peng, G., et al., *Diagnosing lung cancer in exhaled breath using gold nanoparticles.* Nat Nano, 2009. **4**(10): p. 669-673.

16. Moseley, P.T.T.B.C., *Solid-state gas sensors*. 1987, Bristol; Philadelphia: A. Hilger.

17. Semancik, S. and R. Cavicchi, *Kinetically Controlled Chemical Sensing Using Micromachined Structures.* Accounts of Chemical Research, 1998. **31**(5): p. 279-287.

18. Rakow, N.A. and K.S. Suslick, *A colorimetric sensor array for odour visualization.* Nature, 2000. **406**(6797): p. 710-713.

19. Doleman, B.J. and N.S. Lewis, *Comparison of odor detection thresholds and odor discriminablities of a conducting polymer composite electronic nose versus mammalian olfaction.* Sensors and Actuators B: Chemical, 2001. **72**(1): p. 41-50.

20. Lee, S.H., et al., *Mimicking the human smell sensing mechanism with an artificial nose platform.* Biomaterials, 2012. **33**(6): p. 1722-1729.

21. Hagleitner, C., et al., *Smart single-chip gas sensor microsystem.* Nature, 2001. **414**(6861): p. 293-296.

22. Jin, X., et al., *Multichannel Monolithic Quartz Crystal Microbalance Gas Sensor Array.* Analytical Chemistry, 2008. **81**(2): p. 595-603.

23. Martinelli, E., et al., *Short time gas delivery pattern improves long-term sensor reproducibility.* Sensors and Actuators B: Chemical, 2011. **156**(2): p. 753-759.

24. Penza, M., et al., *Tungsten trioxide (WO3) sputtered thin films for a NOx gas sensor.* Sensors and Actuators B: Chemical, 1998. **50**(1): p. 9-18.

25. Raman, B., et al., *Designing and optimizing microsensor arrays for recognizing chemical hazards in complex environments.* Sensors and Actuators B: Chemical, 2009. **137**(2): p. 617-629.

26. Raman, B., et al., *Detecting and recognizing chemical targets in untrained backgrounds with temperature programmed sensors.* Sensors Journal, IEEE, 2012. **12**(11): p. 3238-3247.

27. Hiranaka, Y., T. Abe, and H. Murata, *Gas-dependent response in the temperature transient of SnO2 gas sensors.* Sensors and Actuators B: Chemical, 1992. **9**(3): p. 177-182.

28. Vilanova, X., et al., *Analysis of the conductance transient in thick-film tin oxide gas sensors.* Sensors and Actuators B: Chemical, 1996. **31**(3): p. 175-180.

29. Llobet, E., X. Vilanova, and X. Correig. *Novel technique to identify hazardous gases/vapors based on transient response measurements of tin oxide gas sensors conductance*. in *Proc. SPIE 2504, Environmental Monitoring and Hazardous Waste Site Remediation*. 1995. Munich, Germany.

30. Vergara, A., et al., *Quantitative gas mixture analysis using temperature-modulated micro-hotplate gas sensors: Selection and validation of the optimal modulating frequencies.* Sensors and Actuators B: Chemical, 2007. **123**(2): p. 1002-1016.

31. Distante, C., et al., *On the study of feature extraction methods for an electronic nose.* Sensors and Actuators B: Chemical, 2002. **87**(2): p. 274-288.

32. Zhang, S., et al., *An entire feature extraction method of metal oxide gas sensors.* Sensors and Actuators B: Chemical, 2008. **132**(1): p. 81-89.

33. Llobet, E., et al., *Qualitative and quantitative analysis of volatile organic compounds using transient and steady-state responses of a thick-film tin oxide gas sensor array.* Sensors and Actuators B: Chemical, 1997. **41**(1–3): p. 13-21.

34. Maclay, G.J., J.R. Stetter, and S. Christesen, *Use of time-dependent chemical sensor signals for selective identification.* Sensors and Actuators, 1989. **20**(3): p. 277-285.

35. Niebling, G. and R. Müller, *Non-linear signal evaluation with linear regression techniques for redundant signals.* Sensors and Actuators B: Chemical, 1995. **25**(1–3): p. 805-807.

36. Shafiqul Islam, A.K.M., et al., *Transient parameters of a coated quartz crystal microbalance sensor for the detection of volatile organic compounds (VOCs).* Sensors and Actuators B: Chemical, 2005. **109**(2): p. 238-243.

37. Martelli, C., J.R. Carlson, and T. Emonet, *Intensity Invariant Dynamics and Odor-Specific Latencies in Olfactory Receptor Neuron Response.* The Journal of Neuroscience, 2013. **33**(15): p. 6285-6297.

38. Riffell, J.A., et al., *Flower discrimination by pollinators in a dynamic chemical environment.* Science, 2014. **344**(6191): p. 1515-1518.

39. Stitzel, S.E., D.R. Stein, and D.R. Walt, *Enhancing Vapor Sensor Discrimination by Mimicking a Canine Nasal Cavity Flow Environment.* Journal of the American Chemical Society, 2003. **125**(13): p. 3684-3685.

40. Su, C.-Y., et al., *Temporal coding of odor mixtures in an olfactory receptor neuron.* Proceedings of the National Academy of Sciences, 2011. **108**(12): p. 5075-5080.

41. Vickers, N.J., et al., *Odour-plume dynamics influence the brain's olfactory code.* Nature, 2001. **410**(6827): p. 466-470.

42. Saha, D., et al., *A spatiotemporal coding mechanism for background-invariant odor recognition.* Nat Neurosci, 2013. **16**(12): p. 1830-1839.

43. Mazor, O. and G. Laurent, *Transient Dynamics versus Fixed Points in Odor Representations by Locust Antennal Lobe Projection Neurons.* Neuron, 2005. **48**(4): p. 661-673.

44. Galán, R.F., et al., *Odor-Driven Attractor Dynamics in the Antennal Lobe Allow for Simple and Rapid Olfactory Pattern Classification.* Neural Computation, 2004. **16**(5): p. 999-1012.

45. Artursson, T., et al., *Drift correction for gas sensors using multivariate methods.* Journal of Chemometrics, 2000. **14**(5-6): p. 711-723.

46. Hierlemann, A. and R. Gutierrez-Osuna, *Higher-Order Chemical Sensing.* Chemical Reviews, 2008. **108**(2): p. 563-613.

47. Vergara, A., et al., *Chemical gas sensor drift compensation using classifier ensembles.* Sensors and Actuators B: Chemical, 2012. **166–167**(0): p. 320-329.

48. Zuppa, M., et al., *Drift counteraction with multiple self-organising maps for an electronic nose.* Sensors and Actuators B: Chemical, 2004. **98**(2–3): p. 305-317.

49. Saha, D., et al., *Behavioural correlates of combinatorial versus temporal features of odour codes.* Nat Commun, 2015. **6**.

50. Murlis, J., J.S. Elkinton, and R.T. Cardé, *Odor Plumes and How Insects Use Them.* Annual Review of Entomology, 1992. **37**(1): p. 505-532.

51. Boger, Z., et al., *Rapid Identification of Chemical Warfare Agents by Artificial Neural Network Pruning of Temperature-Programmed Microsensor Databases.* Sensor Letters, 2003. **1**(1): p. 86-92.

52. Sears, W.M., K. Colbow, and F. Consadori, *Algorithms to improve the selectivity of thermally-cycled tin oxide gas sensors.* Sensors and Actuators, 1989. **19**(4): p. 333-349.

53. Li, J., et al., *Carbon Nanotube Sensors for Gas and Organic Vapor Detection.* Nano Letters, 2003. **3**(7): p. 929-933.

54. Ljung, L., *System Identification—Theory for the User, 2nd editionPTR Prentice Hall.* Upper Saddle River, NJ, 1999.

55. Buračas, G.T. and G.M. Boynton, *Efficient Design of Event-Related fMRI Experiments Using M-Sequences.* NeuroImage, 2002. **16**(3, Part A): p. 801-813.

56. Vergara, A., et al., *Optimized temperature modulation of micro-hotplate gas sensors through pseudorandom binary sequences.* Sensors Journal, IEEE, 2005. **5**(6): p. 1369-1378.

57. Ziyatdinov, A., et al., *Bioinspired Early Detection through Gas Flow Modulation in Chemo-Sensory Systems.* Sensors and Actuators B: Chemical, 2015. **206**: p. 538-547.

58. Suzuki, H., *Antennal movements induced by odour and central projection of the antennal neurones in the honey-bee.* Journal of Insect Physiology, 1975. **21**(4): p. 831-847.

59. Laing, D.G., *Identification of single dissimilar odors is achieved by humans with a single sniff.* Physiology & Behavior, 1986. **37**(1): p. 163-170.

60. Woodka, M.D., B.S. Brunschwig, and N.S. Lewis, *Use of Spatiotemporal Response Information from Sorption-Based Sensor Arrays to Identify and Quantify the Composition of Analyte Mixtures.* Langmuir, 2007. **23**(26): p. 13232-13241.

61. Vergara, A., et al., *Demonstration of Fast and Accurate Discrimination and Quantification of Chemically Similar Species Utilizing a Single Cross-Selective Chemiresistor.* Analytical Chemistry, 2014. **86**(14): p. 6753-6757.

62. Cavicchi, R.E., et al., *Fast temperature programmed sensing for micro-hotplate gas sensors.* Electron Device Letters, IEEE, 1995. **16**(6): p. 286-288.

63. Cavicchi, R.E., et al., *Optimized temperature-pulse sequences for the enhancement of chemically specific response patterns from micro-hotplate gas sensors.* Sensors and Actuators B: Chemical, 1996. **33**(1–3): p. 142-146.

64. Gosangi, R. and R. Gutierrez-Osuna, *Active Temperature Programming for Metal-Oxide Chemoresistors.* Sensors Journal, IEEE, 2010. **10**(6): p. 1075-1082.

Chapter 4: Variable compression of intensity in the olfactory circuit

4.1 Introduction

An encounter with an odorant can occur under diverse ambient conditions. For example, environmental variables that are important to chemical sensing such as flow rate, temperature, and humidity can change independently or in combination with one another. Yet, odor detection and recognition must remain invariant to such changes in ambient conditions. How such extrinsic variations (not related to the sensory stimuli of interest) can be filtered out remains to be understood. Here, we examined this issue in the locust olfactory system.

The impact of humidity, in particular, on olfaction is still not understood. Studies report conflicting impacts on the perceptual effects of humidity variations. Some find that limits of detection are lowered with increases in humidity for humans [1] and other animal models. Others, however, have found no significant changes in perception under varying RH conditions [2, 3]. How do changes in relative humidity (RH) influence odor recognition?

Detection of humidity is important for animals for survival, reproduction, and comfort [4-8]. Insects, including fruit flies (*D. melanogaster*) [9] and locusts [10, 11], have specialized sensors, olfactory receptors neurons (ORNs), for detecting changes in humidity levels. Hygrosensory ORNs come in two varieties, those that detect dry conditions and those moist conditions. In locusts, they are housed in a type of sensory hair (*Sensillum*) on the insect antenna referred to as *coeloconica*.

Changes in the relative humidity (RH) levels can also be thought of as an interfering stimulus on top of which olfactory cues have to be detected. Biological sensory systems have developed robust processing techniques to filter out background or interfering stimuli. Previous work by us and others[12, 13] have revealed that changing backgrounds may initially perturb a sensory response, yet recognition will remain invariant as long as the same set of neurons, or significant subset thereof, is activated. Whether this is the case and the same circuit mechanisms that allows background invariant recognition can also facilitate humidity invariance is not clear.

In this study, we explored how humidity affects olfactory signal detection and processing. Our results reveal that at the ORN level, changes in RH significantly altered the net spontaneous activity of the population, but only had minimal effect on the stimulus-induced response magnitude. However, this variation in the overall sensory input triggered the following antennal lobe circuits to process sensory inputs elicited by the same stimulus in humidity-dependent fashion. Our results indicate that the odor identity was robustly encoded irrespective of the relative humidity level. However, we found that information regarding stimulus intensity is compressed at high relative humidity conditions. This alteration in the neural response dynamic range resulted in predictable changes in the behavioral responses of locusts to conditioned stimulus presented at different intensities. Therefore, our results reveal that olfactory systems robustly encode odorant identity and concentration in addition to the effects of environmental variability and this information has an effect on behavioral outcomes.

4.2 Results

4.2.1 Humidity does not affect sensory responses to stimuli

We began by examining how changes in relative humidity levels altered the sensory input from ORNs. A simple way to measure bulk ORN signal is by monitoring the potential difference across the two ends of the antenna, in a recording known as an electroanntenogram (EAG) [14]. Increases in EAG signal represent an increase in population ORN activity while decreases reflect a decrease. To reduce the effects of other environmental variations and improve the stability of our EAG recordings, we employed a novel EAG method that allowed for use of an intact locust prep, yet isolated the antenna from external environmental variations (Figure 4.1 B).

We measured the EAG responses to 3 concentrations of hexanol delivered in four relative humidity levels, 0%, 33%, 66%, and 100% RH. As expected, we observed that the magnitude of stimulus-evoked EAG response monotonically varied with the concentration of odorant used. The EAG magnitude to hexanol 0.1% was found to be significantly different than the other two concentrations tested (Student's t-test, $P < 0.05$ with Bonferroni correction). There was not a significant difference in the response magnitude between 1% or 10% hexanol at 0% and 100% RH (Figure 4.1, C-D). We found that the magnitude of response was largely invariant to the background humidity level (Figure 4.1E). However, there was a small portion of cases in which there was a significant difference between the 0% and 100% RH.

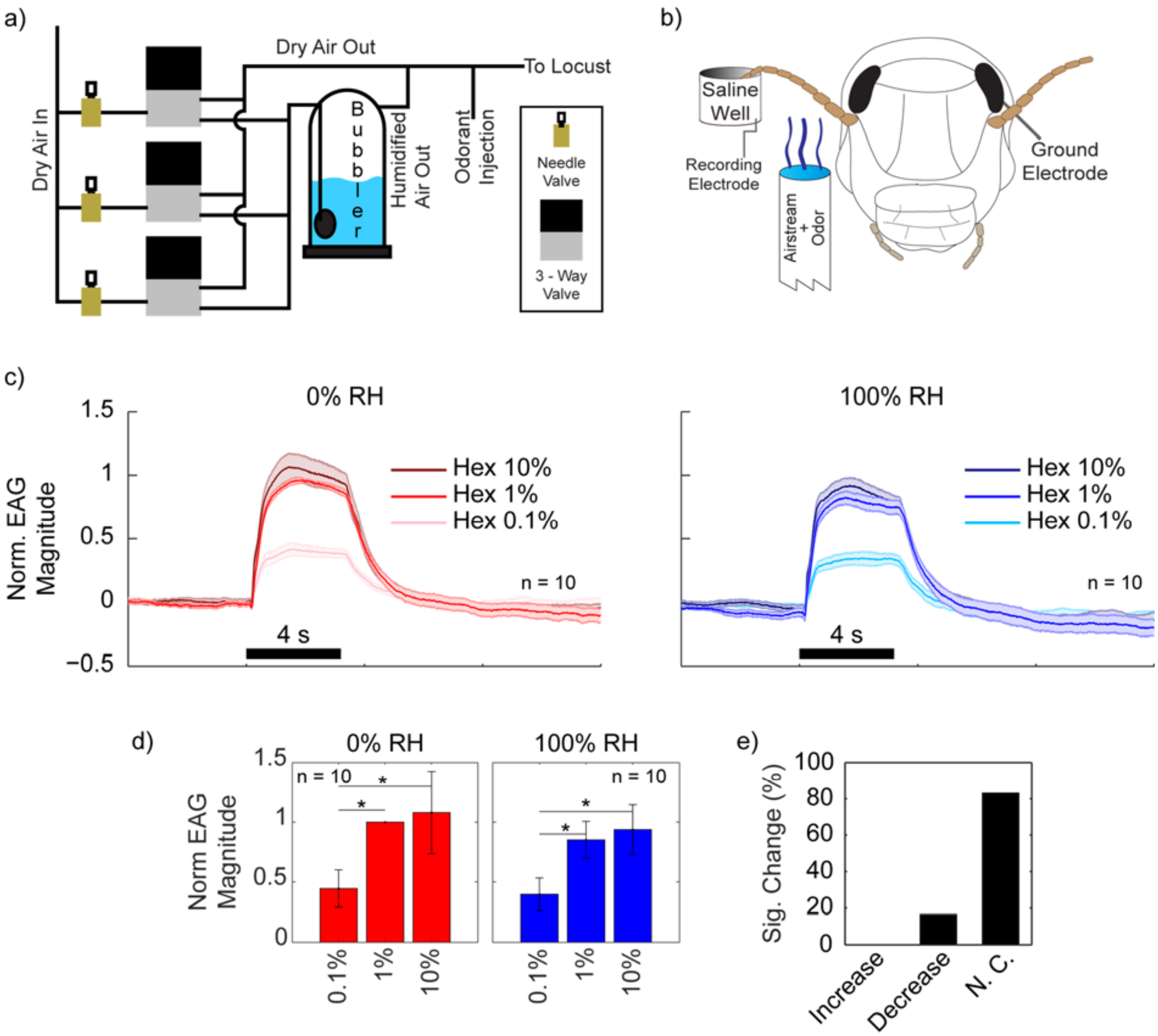
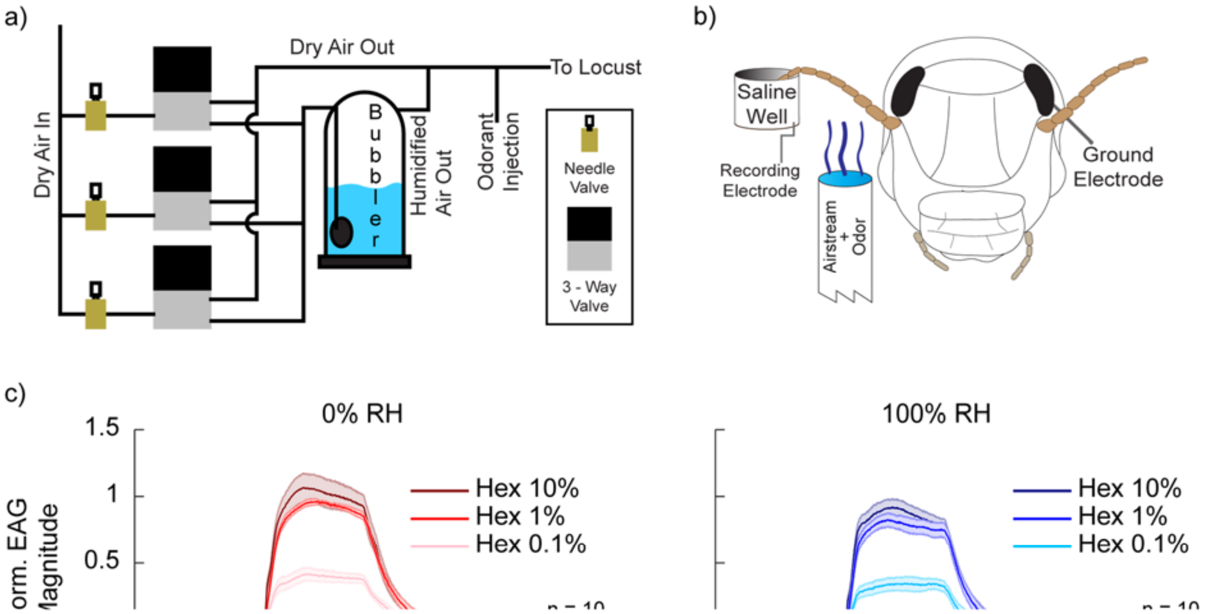


Figure 4.1. Humidification apparatus and Locust EAGs. a) Clean, dehumidified air flows through a valve system. Valves deliver air directly to the locust or through a bubbler for humidification before mixing with dry air delivered to the locust. The ratio of dry air vs humidified air controls overall humidity level of the airstream to the locust. b) Schematic drawing of locust EAG. Dorsal tip of an antenna is cut and dipped in a well of Locust Ringer's solution. A recording electrode is inserted in the well and ground electrode is inserted in the contra-lateral eye. c) Representative EAGs showing potential difference when locust is exposed to 0% RH and 100% RH carrier air. Signals are baseline aligned and normalized to the mean Hexanol 1%, 0% RH response. Shaded bands indicate S.E.M. d) Differences in EAG response magnitudes as a function of concentration in both 0% and 100% RH conditions. In both RH levels, 0.1% hexanol has a significantly different response magnitude (Two-tailed t-test, $P < 0.05$, Bonferroni corrected) than the higher two concentrations. The difference between the 1% and 10% responses was not found to be significant in either humidity level. e) Summary statistics of EAG magnitude change between 0% and 100% RH for 30 locust-odorant pairs. Significant differences are calculated with a one tailed t-test ($P < 0.05$, Bonferroni corrected). A 100% RH carrier stream did not cause

significant changes in the stimulus evoked response magnitude.

4.2.2 Humidity alters the total ORN baseline activity

Surprisingly, we found that increasing humidity levels changed the baseline activity of the EAGs (Figure 4.2). Our results show that in 63% of locust-odorant pairs, there was a significant decrease in baseline activity from 0% RH to 100% RH (one tailed t-test, $p < 0.05$, Bonferroni corrected). Furthermore, this result was robust as humidity was cycled from 0% RH to 100% RH and back.



4.2.3 Odorant identity is encoded independently of humidity levels

To examine whether information about the RH influences downstream processing, we performed extracellular recordings of projection neurons (PNs) in the locust antennal lobe. We found that changing humidity levels did not affect odorant specificity of PNs (Figure 4.2A-B). Consistent with previously published results, we found that different odorants elicited different responses in neurons [12, 15-18].

We found that changes in humidity level generally did not affect a neuron's responsiveness to an odorant. We found that most neurons had correlations of 0.5 or higher for neural activity in 0% and 100% RH conditions across odorants tested (Figure 4.3 B). Therefore, PNs that were responsive stayed responsive while those that were not responsive remained non-responsive.

To assess how the ensemble antennal lobe response changed to different stimuli under different conditions, the total spike count during the stimulus (4 s, mean of 5 trials) was correlated across odorants and RH levels (Figure 4.3 B-C). Correlations were found to be low across different odorants indicating that different sets of neurons were activated by different stimuli. Interestingly, the off-diagonals, comparing responses to the same odorant in 0% and 100% RH showed high levels of correlation, thus, indicating that a conserved set of neurons was activated across humidity conditions.

To gain a better understanding of the changes occurring in the antennal lobe, we performed a qualitative dimensionality reduction using principal component analysis (PCA) to visualize the activity across all the PNs. The activity of 82 PNs was condensed to 3 dimensions which captured 33.23% of the variance (Figure 4.3 D). Note that different odorants elicited trajectories that evolved in different directions after dimensionality reduction. However, variation in RH did not alter the direction of the trajectory for any odorant. In sum, the correlation and dimensionality reduction analyses suggest that the stimulus identity is robustly maintained across different humidity conditions.

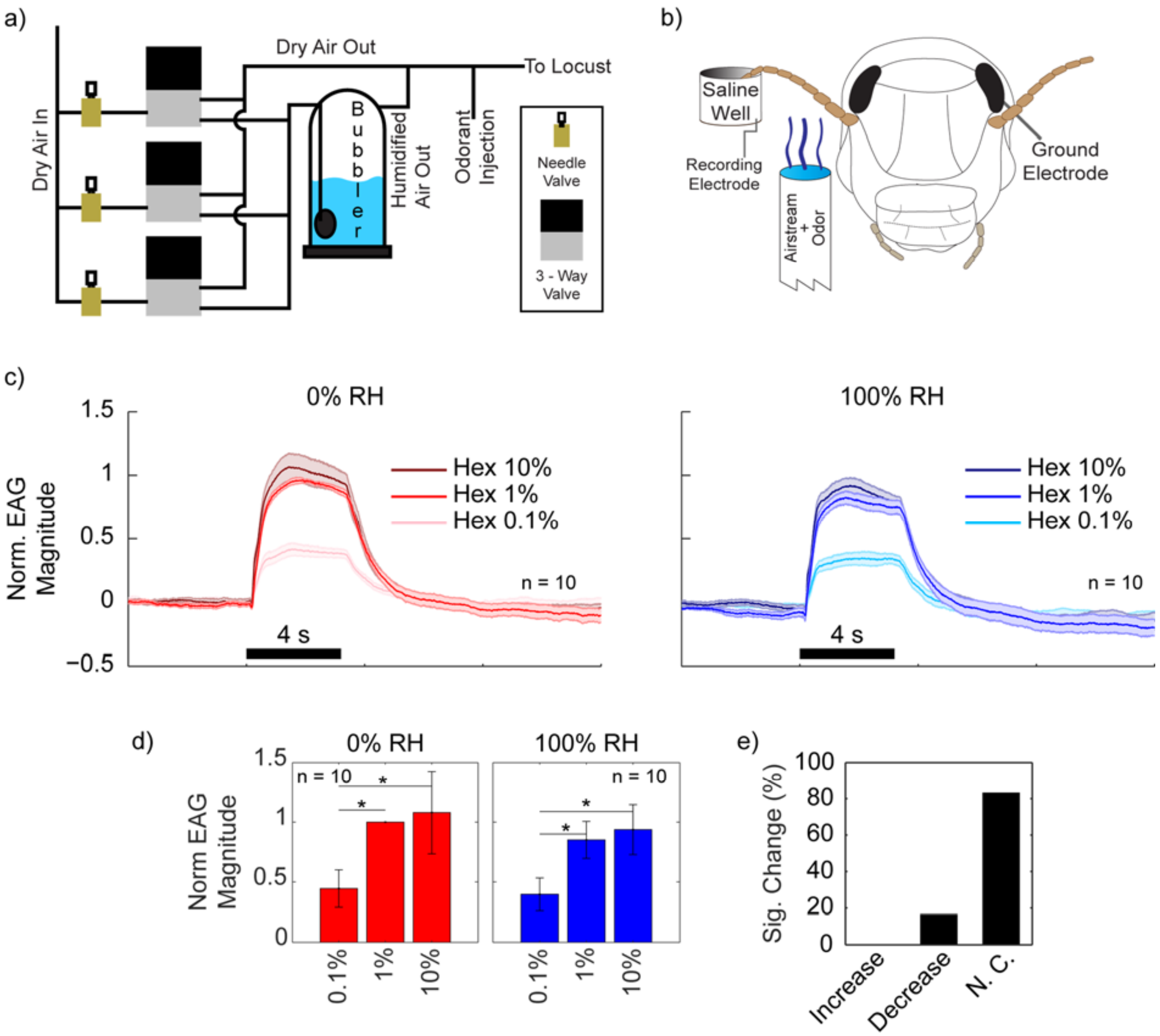
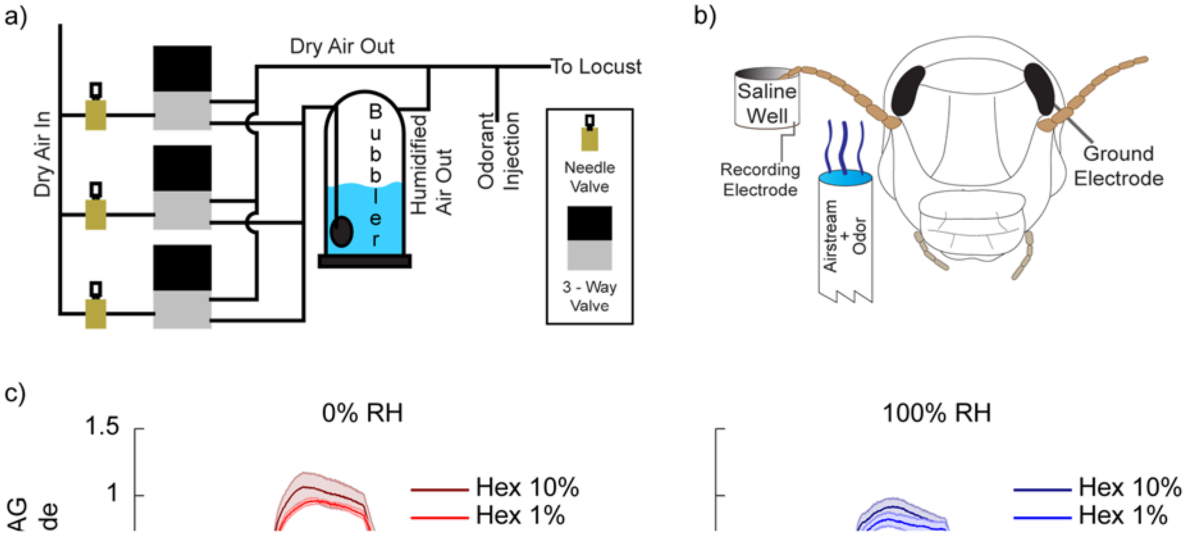


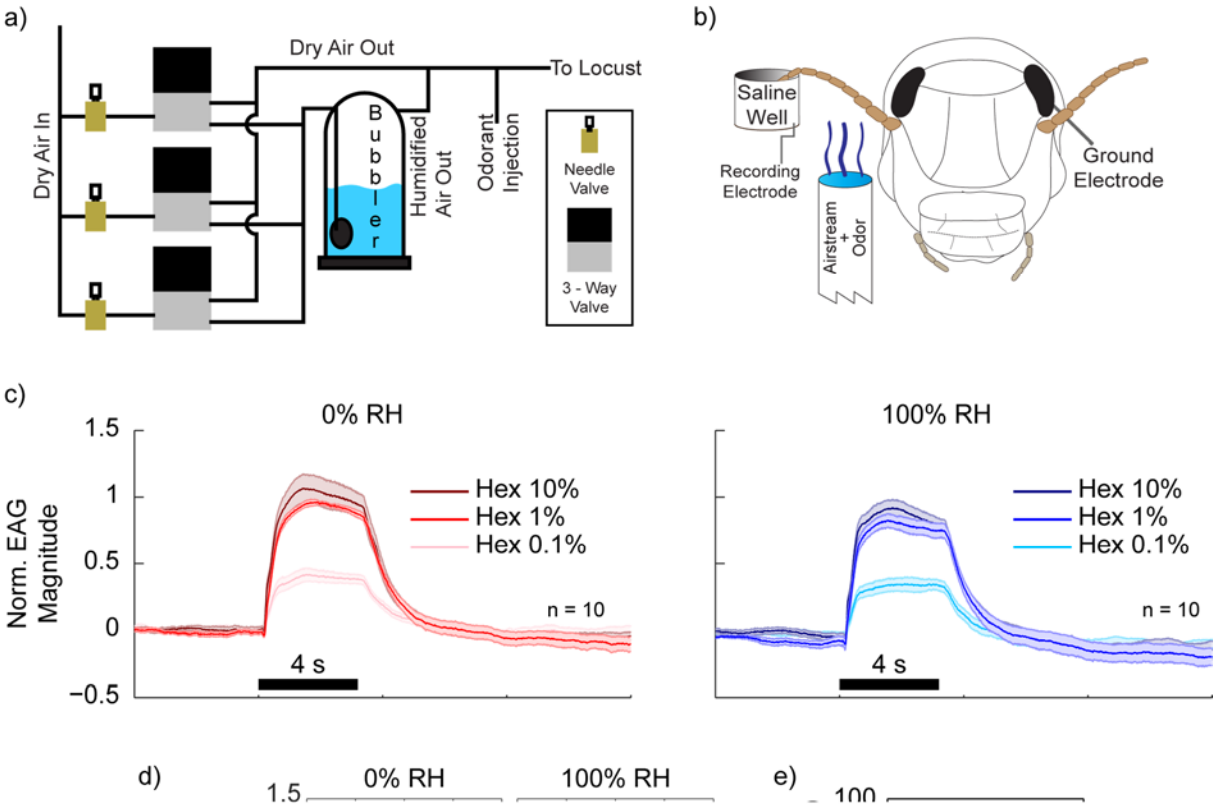
Figure 4.3. Odorant identity is preserved across RH changes. A) PN tuning does not change with RH changes. Responses of 4 PNs to 4 odorants (Hexanol, 1-Octanol, Isoamyl Acetate, and Cyclohexanone, all 1% (v/v)). Rasters show 10 trials for each odorant: top 5 (red background) are in 0% RH, bottom 5 (blue background) are in 100% RH. B) Probability density function histograms for correlations of 10 second (2s pre-stimulus, 4s stimulus, 4s post-stimulus) projection neuron PSTHs in 0% and 100% RH environments. C) Correlogram of total neural activity during 4 second stimulus. Correlations are higher for the same odorant across RH levels than across odorants. D) Principle component analysis (PCA) of 4 odorants in 0% RH (solid lines) and 100% RH (dashed lines) during 4 second odor evoked response. Trajectories for odorants lie in the same sub-space indicating odorant identity is conserved across humidity levels.

4.2.4 Odorant intensity is compressed at high humidity conditions

As stimulus intensity decreases, typically fewer neurons respond (Figure 4.4) and the activity in many responding neurons reduces, thereby resulting in a reduction in the total population activity (i.e. decreasing the trajectory length). We found that for most odorants the length of the PN response trajectories decreased when the same odorant was presented in 100% RH conditions (Figure 4.5). This trend was observed for all odorants tested except IAA, for which there was little to no change in trajectory length across humidity extremes. It has been previously reported that when stimulus intensity is changed, the trajectory length correspondingly changes [18]. Therefore, we next examined whether changes in RH altered how stimulus intensity was encoded.



Does the same mechanism also cause reduction in ensemble PN activity across the two humidity conditions? To determine this, we examined the PN activity in response to a hexanol stimulus delivered at 3 concentrations in both 0% and 100% RH conditions. At 0% RH, as the odor concentration decreased, PNs tended to respond more weakly and the stimulus recruited fewer neurons (Figure 4.6 A). The trend is also evident in raster plots and trial-averaged PSTHs showing the responses of single neurons (Figure 4.6 B). The responses to the same stimuli was presented at 100% RH conditions are shown in the right panels for comparison. Two observations are noteworthy: (i) the response of the PN at the high concentration reduces. (ii) the response of the PNs at the lower concentration increases.



Population PSTHs averaging across all 94 PNs in the dataset also show concentration dependent responses (Figure 4.6 C-D). The difference in response peaks, however, are compressed compared to the 0% RH conditions. Note that in 100% RH conditions there are still subtle decreases with decreasing concentration. Interestingly, no noticeable change in response latency, post-stimulus activity, or spontaneous pre-stimulus activity were observed. The latter is surprising given the changes in the bulk ORN signals that showed a significant change in the baseline response across 0% and 100% RH conditions.

Next, we performed a PCA dimensionality reduction analysis to visualize the PN ensemble neural activity across concentrations and at different relative humidity conditions. Consistent with previously reported results [18], as concentration decreases, the trajectory length decreases, but the direction of evolution remained the same. At the higher humidity conditions, we noticed that the differences between the length and span of the trajectories across concentrations decreased noticeably (Figure 4.6 E-F). Interestingly, the length of the trajectories notably decreased at higher intensities, while there was a modest increase in the low intensity responses. Our results taken together indicate that the compression of intensity information can vary depending on the ambient conditions.

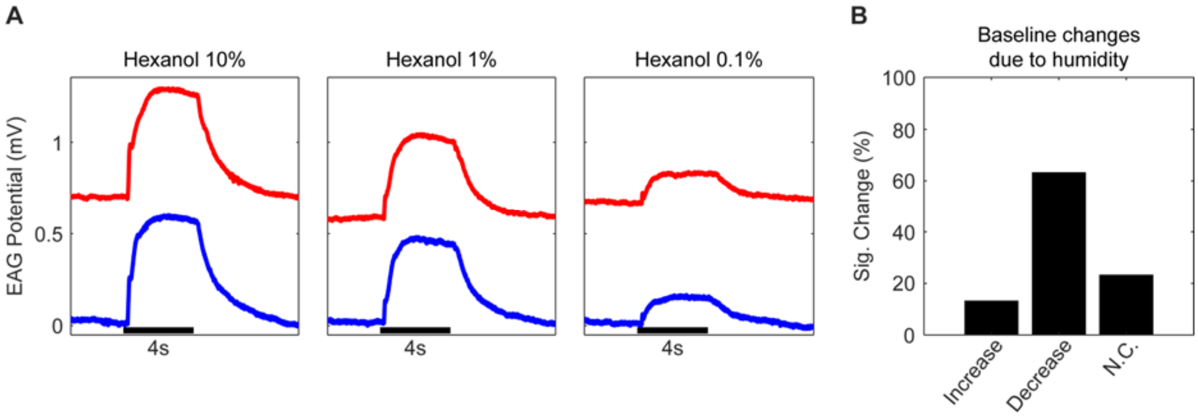


Figure 4.6. Projection neurons have concentration dependent responses in dry conditions but not in humid conditions. a1-3) Raster plots of PNs responding to 3 concentrations of hexanol, 0.1%, 1%, and 10% (v/v) with 5 trials for each concentration. Shaded color bar indicates stimulus delivery duration of 4 seconds. 0% RH conditions are shown on the left for each PN and 100% RH shown on the right for each PN. b1-3) Peri-stimulus time histograms corresponding to PNs in a1-3. Three traces are shown, 1 for each concentration delivered. 0% RH conditions are shown on the left for each PN and 100% RH shown on the right for each PN. c) Population PSTHs of 94 PNs in 0% RH. One trace is shown for each concentration delivered. Colored circles indicate peak response magnitudes and shaded regions indicate S.E.M. d) Same as c, but for 100% RH. e) PN response trajectories in 0% RH. Over 27% of the variance is captured in the first two principle components. f) PN response trajectories in 100% RH.

4.2.5 Decoding stimulus intensities in different ambient conditions

Increases in RH decrease neural activity in a manner similar to stimulus intensity changes. Yet, it is unclear whether the mechanisms behind both phenomena are the same. To better understand and quantify the changes in the stimulus-evoked activity, we compared the changes in neural activity across stimulus intensity (Figure 4.7 A). We found that the ensemble response correlations are higher within an intensity across RH levels than they are across stimulus intensities (Figure 4.7 B, off-diagonals). Furthermore, the correlations were more similar between hexanol 1% and 10% than it was for hexanol 1% and 0.1%. Therefore, although the firing rates of many individual PNs change with humidity level, our results indicate that the information about stimulus intensity is preserved as combinations of activated PNs are conserved across conditions. These observations were validated using a quantitative classification analysis using the high-dimensional PN response vectors.

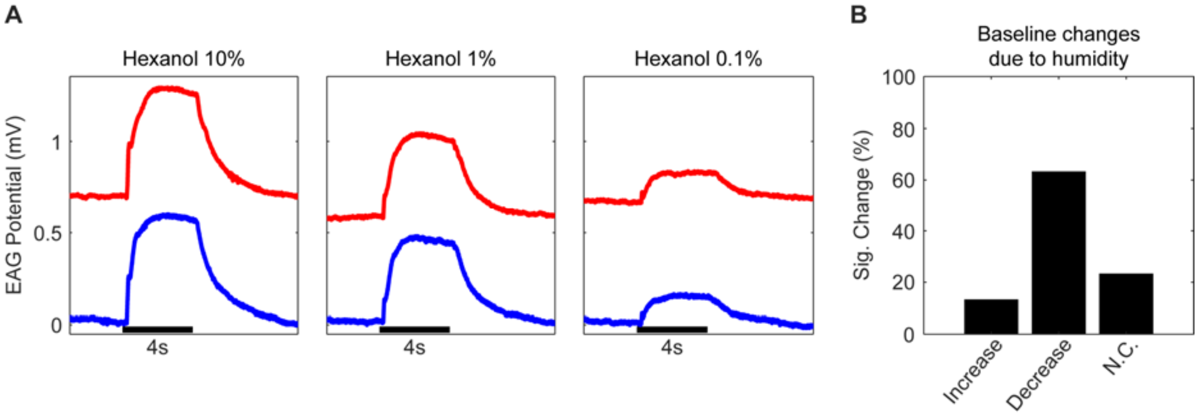


Figure 4.7. Different populations of neurons respond to presence of humidity effects than changes in concentration. a) Mean firing rate for population of PNs during a Hexanol stimulus. Neurons are sorted in descending order based on firing rate a 1% Hexanol stimulus in 0% RH (red trace). Neural variability for other concentrations (0.1% and 10%; pink and brown traces, respectively) in 0% RH is higher than in 1% Hexanol in 100% RH (blue trace). b) Correlogram of total neural activity during 4 second stimulus across intensities and humidity levels. c) Correlations of stimulus evoked response of different conditions to 1% Hexanol in 0% RH. Correlation values decrease primarily as a function of concentration. Note, that stimuli in 100% RH conditions have lower correlations than comparable stimuli in 0% RH. d) Bin-by-bin, high dimensionality classification shows concentration identity is preserved. Rows represent trials and each tick represents a 50ms time bin, 9 seconds of each trial are shown. Lines under each subplot reflect the stimulus duration and the color represents the concentration. All stimuli delivered in 0% RH. e) Probability of classification for each concentration over the 4 second stimulus duration. f) Bar plots reflecting area under the curve in panel d. g-i) Same as d-f but for stimuli delivered in 100% RH.

4.2.6 Neural response variations correlate with changes in behavioral output

Finally, to assess the effect of the observed changes in the neural activity, we performed a behavioral analysis using a previously reported appetitive conditioning paradigm [12, 19]. We trained locusts in a 0% RH (n = 20 locusts) background (Figure 4.8 A). Training was done using 1% hexanol (v/v) as the conditioned stimulus and organic wheat grass as the unconditioned stimulus. Testing was performed under both 0% RH and 100% RH to determine whether humidity levels had an effect on behavioral performance as measured by the palp opening response (POR). We also tested the performance to the three concentrations of hexanol used in our electrophysiology experiments. The order of concentrations used was pseudo-random for each experimental day.

We found that locusts responded robustly to the trained odorant irrespective of the humidity conditions. There was a small but significant reduction in the POR responses at the lowest intensity (hexanol 0.1%) in the 0% RH case (Figure 4.8 B). However, this difference was largely eliminated when the same locusts were tested in 100% RH conditions and there were no significant differences between POR magnitudes (t-test, $P > 0.05$, Bonferroni corrected). These results are in direct agreement with our physiological findings that the concentration information in compressed at high relative humidity conditions.

We examined how different features of the neural activity in the antennal lobe (the magnitude of the population response (‘coarse feature’) as well as similarity with the conditioned odor’s population vector (‘fine feature’)) correlated with the observed behavioral responses. We found that both the coarse and fine neural response features had good linear fits with POR magnitudes (Figure 4.8 C). The correlations to neural activity, however, were better predictors of peak POR magnitude (Rcorr vs. Rmag = 0.83 vs 0.69). Taken together, these results indicate that subtle variations in neural response, when present, translate to corresponding changes in behavioral responses.

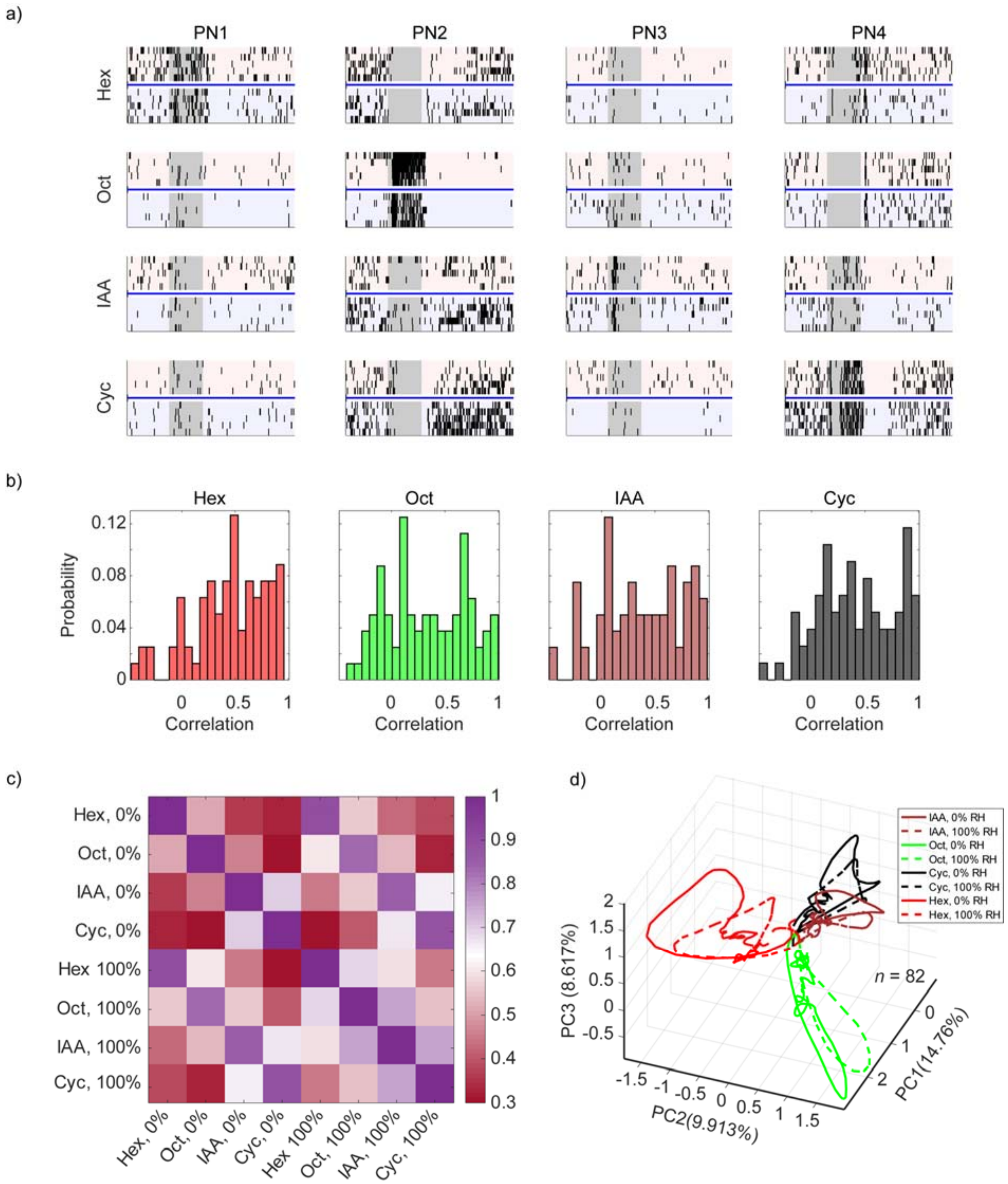


Figure 4.8. Palp opening response is invariant to RH levels. a) Schematic of Locust POR recordings. PORs are measured as distance between the distal segments of two palps (painted green). b) PORs for 3 concentrations of Hexanol in 00RH and 100RH conditions. There is no significant difference between PORs in RH levels. Locusts were trained with Hexanol 1% in 00 RH. Shaded bands indicate S.E.M. c-d) Comparison of peak POR to peak PSTH magnitude (c) and peak POR to neural activity correlations (d). PORs from locusts trained

under 00RH conditions. Linear fit indicates that neural activity correlations are a stronger indicator of POR performance than PSTH magnitudes.

4.3 Discussion

The olfactory system is robust to changes in the environment. These changes include, but are not limited to changes in flow rate, temperature, background volatiles present, light, circadian rhythms, and many others. In olfactory neuroscience, these parameters are generally held constant as controls. Olfactory systems, however, are able to perform under constantly varying external conditions. Here we have examined one aspect of environmental variability, relative humidity, and determined the mechanisms by which the olfactory system is able to cope or adapt to changes in those levels. We observed that the system is able to not only identify an odorant under different humidity conditions, but also retains information about intensity.

There are extrinsic and intrinsic parameters that can be used to alter the activity of the neural system. In this study we can think of modulating RH as a means to alter the intrinsic activity of the system. RH is relatively slow to change and is a persistent effector on the olfactory system. Changes in the RH change the dynamics under which the system is operating and the underlying set point from which other stimuli will be perceived. An odorant stimulus, on the other hand, is an extrinsic parameter that also modulates the activity in the system, but is transient and results in a deviation from the systems set point which is restored soon after the stimulus's removal. Here we found that increases in the environmental RH tend to decrease stimulus evoked activity, but do not alter the neural code enough to change odorant identity information.

The full effects of changes in environmental humidity remain to be assessed. We limited our study to the sensor level and primary processing stage, the antenna and antennal lobe, respectively. Further studies need to be performed to determine how the changes observed in the antennal lobe are translated downstream in the mushroom bodies and their Kenyon cells. We would also want to assess how changes in RH effect local field potentials (LFP) in the mushroom bodies.

At the sensor level, this study focused on the effects of RH on EAG magnitude. We also found that there was an effect of RH change on EAG baseline. Trends showed that EAG baseline potentials decreased with RH increases (Figure 4.2). This indicates that overall activity of ORNs decrease with RH increases and that a weaker input is delivered to the antennal lobe circuits. These trends proved to be too unreliable to report in this study but are a strong candidate for future studies. We also attempted to collect data for an even lower concentration, Hexanol 0.01%. In our attempts, we found that the odorant source was quickly depleted thus providing unreliable data, and not feasible to accomplish with our experimental methods.

We found that humidity did not cause compression in the trajectories for isoamyl acetate 1% (Figure 4.5), unlike the other three odorants tested. There are a few possible explanations for why this may have occurred. First, a 1% concentration of IAA maybe too low to evoke this effect. Our results for Hexanol 0.1% also did not show compression; IAA may have the same phenomenon at higher concentrations. Second, IAA, has been thought to be a pheromone for locusts and maybe sensed by different sensilla (*sensilla trichodea*). It may be that this class of sensilla is affected differently by water vapor than the primary sensilla for the other odorants (*sensilla basiconica*).

In our behavioral experiments we also trained locusts under 100% RH conditions. The PORs for these locusts were quite different than those trained in 0% RH (Supplemental Figure 5). Two aspects stand out as particularly distinct. The first is that concentration dependence is restored for PORs in 0% RH. The second is that when these locusts are tested in 100% RH conditions, compression of the response across concentrations is no longer observed. Instead, we see that the POR for Hexanol 1% and 10% are actually larger than in the dry cases and their peaks are significantly ($p < 0.05$, Bonferroni corrected) higher than the peak of Hexanol 0.1%. Interestingly, we see the same relation between the Peak POR and neural information from the peak PSTH and neural correlation with the training stimulus. Here we also see, through linear fitting, that while both parameters are related to the behavioral response, the neural correlation has a stronger relation to the peak POR.

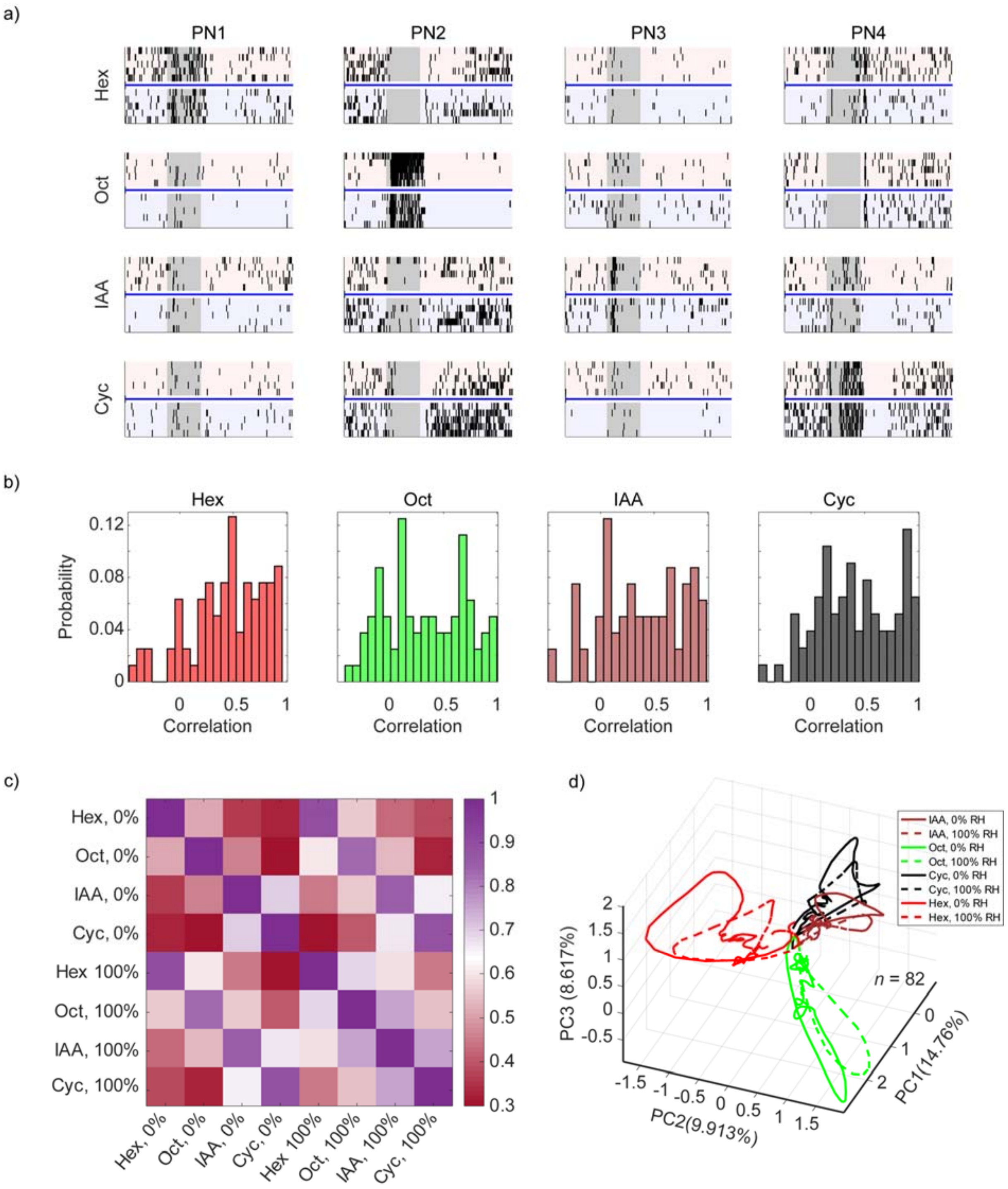


Figure 4.9. Palp opening response is stronger in locusts trained in 100% RH. a) Schematic of Locust POR recordings. PORs are measured as distance between the distal segments of two palps (painted green). b) PORs for 3 concentrations of Hexanol in 00RH and 100RH conditions. There is no significant difference between PORs in RH levels. Locusts were

trained with Hexanol 1% in 00 RH. Shaded bands indicate S.E.M. c-d) Comparison of peak POR to peak PSTH magnitude (c) and peak POR to neural activity correlations (d). PORs from locusts trained under 00RH conditions. Linear fit indicates that neural activity correlations are a stronger indicator of POR performance than PSTH magnitudes.

This study has explored the effects of variations in environmental conditions on olfactory coding and behavioral output in a simple olfactory system. We found that increases in RH cause compression of the neural activity in the antennal lobe, which is not seen at the upstream antenna. Furthermore, we found that the mechanism by which this compression occurs is different than that which occurs for different stimulus intensities. This suggests that there two different mechanisms by which the neural responses can be perturbed: 1) by changing the stimulus intensity or 2) by changing the environment in which the stimulus is delivered. These changes affect different sub-populations of neurons but have similar results. The neural activity changes but odorant identity is preserved. Our findings show that RH can be used to alter the intrinsic operation of the system and put it in a different regime opening further avenues to test the operation of the olfactory system and its edge cases.

4.4 Author contributions

The work in this chapter is part of the manuscript entitled: Katta, N., Zhang, L., Chandak, R., O’Neill, M. and Raman, B. (2017). "Variable compression of stimulus intensity in an olfactory circuit."

Experiments were designed by NK, MO, and BR. EAG and PN data was collected by NK. Behavior data was collected by RC. Data analysis was performed by NK, LZ, and RC.

4.5 References

1. Kuehn, M., et al., *Changes of pressure and humidity affect olfactory function.* European Archives of Oto-Rhino-Laryngology, 2008. **265**(3): p. 299-302.

2. Philpott, C., et al., *The effect of temperature, humidity and peak inspiratory nasal flow on olfactory thresholds.* Clinical Otolaryngology & Allied Sciences, 2004. **29**(1): p. 24-31.

3. Alfonso Collado, J.I. and H. Vallés Varela, *Study of Olfactory Function for Pyridine in Healthy Population: Influence of Variations in Humidity.* Acta Otorrinolaringologica (English Edition), 2008. **59**(10): p. 475-479.

4. Shelford, V.E., *A COMPARISON OF THE RESPONSES OF ANIMALS IN GRADIENTS OF ENVIRONMENTAL FACTORS WITH PARTICULAR REFERENCE TO THE METHOD OF REACTION OF REPRESENTATIVES OF THE VARIOUS GROUPS FROM PROTOZOA TO MAMMALS.* Science, 1918. **48**(1235): p. 225-230.

5. Montell, C., *TRP Channels: It's Not the Heat, It's the Humidity.* Current Biology, 2008. **18**(3): p. R123-R126.

6. Liu, L., et al., *Drosophila hygrosensation requires the TRP channels water witch and nanchung.* Nature, 2007. **450**(7167): p. 294-298.

7. Strusberg, I., et al., *Influence of weather conditions on rheumatic pain.* The Journal of Rheumatology, 2002. **29**(2): p. 335-338.

8. Krång, A.-S., et al., *Transition from sea to land: olfactory function and constraints in the terrestrial hermit crab Coenobita clypeatus.* Proceedings of the Royal Society B: Biological Sciences, 2012. **279**(1742): p. 3510-3519.

9. Enjin, A., et al., *Humidity Sensing in Drosophila.* Current Biology, 2016. **26**(10): p. 1352-1358.

10. Altner, H., C. Routil, and R. Loftus, *The structure of bimodal chemo-, thermo-, and hygroreceptive sensilla on the antenna of Locusta migratoria.* Cell and Tissue Research, 1981. **215**(2): p. 289-308.

11. Ochieng', S.A., E. Hallberg, and B.S. Hansson, *Fine structure and distribution of antennal sensilla of the desert locust, Schistocerca gregaria (Orthoptera: Acrididae).* Cell and Tissue Research, 1998. **291**(3): p. 525-536.

12. Saha, D., et al., *A spatiotemporal coding mechanism for background-invariant odor recognition.* Nat Neurosci, 2013. **16**(12): p. 1830-1839.

13. Rokni, D., et al., *An olfactory cocktail party: figure-ground segregation of odorants in rodents.* Nat Neurosci, 2014. **17**(9): p. 1225-1232.

14. Roelofs, W., et al., *Sex Attractant of the Codling Moth: Characterization with Electroantennogram Technique.* Science, 1971. **174**(4006): p. 297-299.

15. Raman, B., et al., *Temporally Diverse Firing Patterns in Olfactory Receptor Neurons Underlie Spatiotemporal Neural Codes for Odors.* The Journal of Neuroscience, 2010. **30**(6): p. 1994-2006.

16. Saha, D., et al., *Behavioural correlates of combinatorial versus temporal features of odour codes.* Nat Commun, 2015. **6**.

17. Mazor, O. and G. Laurent, *Transient Dynamics versus Fixed Points in Odor Representations by Locust Antennal Lobe Projection Neurons.* Neuron, 2005. **48**(4): p. 661-673.

18. Stopfer, M., V. Jayaraman, and G. Laurent, *Intensity versus Identity Coding in an Olfactory System.* Neuron, 2003. **39**(6): p. 991-1004.

19. Simões, P., S.R. Ott, and J.E. Niven, *Associative olfactory learning in the desert locust, Schistocerca gregaria.* The Journal of Experimental Biology, 2011. **214**(15): p. 2495-2503.

Chapter 5: Developing a non-invasive chemical sensing approach for detection of malaria biomarkers

In the previous chapter, we considered how non-invasive chemical sensing approaches (both artificial and biological) can be made robust against sensor degradation and variations in environmental factors. Here, I consider a potential target application where a robust chemical sensing solution can be applied. First, I will discuss my recent collaborative work[1] identifying breath biomarkers of malaria in exhaled breath. In particular, I will explain in detail my contributions to this work in developing an analytical approach for the identification of disease biomarkers, in exhaled breath, based on GC/MS datasets.

5.1 Breathprinting malaria

5.1.1 The malaria parasite: Prevalence, symptoms, and diagnosis

Malaria is a disease caused by infection of erythrocytes by parasites in the *Plasmodium* genus. The common symptoms of malaria are fever, headache, chills, and vomiting, which can often be confused with symptoms of other diseases common to regions with malaria. There are currently 91 countries and territories with endemic malaria. In 2015, there were an estimated 212 million cases of infection, which resulted in 429,000 deaths. Of these, 70% were of children under the age of 5 [1]. The African continent accounted for 90% of the infections and 92% of the deaths from malaria in 2015 [2].

---

[1] Schabler et al. Breathprinting Reveals Malaria-Associated Biomarkers and Mosquito Attractants, 2017 (in review). NK contributions: developed an analytical approach for identifying malaria biomarkers in an unbiased fashion. This contribution is detailed in this chapter.

Globally, there are five main species responsible for malaria: *P. falciparum, P. vivax, P. ovale, P. malariae, and P. knowlesi* [3]. The parasite is most commonly spread by bites from infected females of the *Anopheles* mosquito. There are two dominant species of the malaria parasite: *P. falciparum*, which is found mostly in the African continent, and *P. vivax* which is found primarily in Southeast Asia [2].

P. vivax is an especially potent form of the parasite. Reports as early as 1989 suggested *P. vivax* resistance to chloroquine [4], with more recent evidence of spread of resistant strains [5]. Furthermore, radical cure is difficult for this form because of its dormant liver stage [1, 3] which often causes relapses after treatment [5]. This strain, however, is not prevalent in Africa because of the Duffy-negative phenotype, present in central and west Africa, which prevents invasion of the parasite into erythrocytes [3]. Growth of *P. vivax* is particularly seen in regions were the spread of *P. falciparum* has been controlled, such as Southeast Asia.

Despite the success of methods to control *P. falciparum* spread in certain regions, it remains the dominant source of malaria infection and deaths globally. It is primarily endemic to sub-Saharan Africa, where it is responsible for most of the world's malaria-related death and disability [2, 3]. While it is a dominant cause of child death, distinguishing malaria from other diseases remains difficult. The primary symptoms overlap with other illnesses, such as severe sepsis [3]. The parasite sequesters itself in the body by causing infected red blood cells to attach to the lining of small blood vessels, which can restrict blood flow and correlates to with lactic acidosis, particularly in children [6]. Although recent efforts have had excellent results reducing the prevalence and effects of malaria, there remain many challenges. For example, reports indicate the evolution of an artemisinin resistant strain of *P. falciparum* in Southeast Asia [7].

There are a number of different diagnostic methods for malaria, which vary based on location and the available standard of healthcare. The gold standards for diagnosis are thick and thin blood smears [8]. Typical detection limits for trained microscopists are on the order of 100 parasites per 100 μL of blood, the lower limit of pyrogenic density [9]. Lower parasitaemia is more likely to be incidental to another illness, so tests with lower detection limits have higher false positives [9]. Other methods include detection of parasite antigens or nucleic acids. These are commonly referred to a rapid diagnostic tests (RDTs) and are becoming more prevalent worldwide [1, 8, 9]. While both methods have respective advantages, they also come with tradeoffs. Blood smears require maintenance of laboratory equipment as well as training and supervision of staff to ensure proper diagnoses. RDTs, while cheaper and easier to use, cannot detect all strains of malaria, cannot distinguish new infections from recently treated ones, and are not suitable in all regions because of variations in antigen production [9]. Furthermore, RDTs cannot be used to evaluate treatment response and can also lead to false positives due to the presence of residual antigens [8]. Thus, there is a need for a specific, sensitive malaria test which does not require high levels of maintenance or training to perform. Here, I will explore the potential of an alternate diagnostic methodology using exhaled breath to diagnose malaria.

5.1.2 Breath VOC analysis

To identify breath biomarkers, we used a GC/MS dataset provided by our collaborators[2]. Details about sampling collection and processing are discussed in the methods (Sec. 2.4.1-2.4.2) Briefly, breath samples from pediatric patients (children 3-15 years age) at two ambulatory pediatric centers in Lilongwe, Malawi, were included in this dataset. Children requiring urgent

---

[2] Samples were collected by Dr. Indi Trehan and Lucy Bollinger and processed by Chad Schaber and Dr. Audrey R. John Odom. Data to be published in Schaber et al., 201710. Schaber, C., et al., *Breathpriting Reveals Malaria-Associated Biomarkers and Mosquito Attractants.* Manuscript in preparation, 2017..

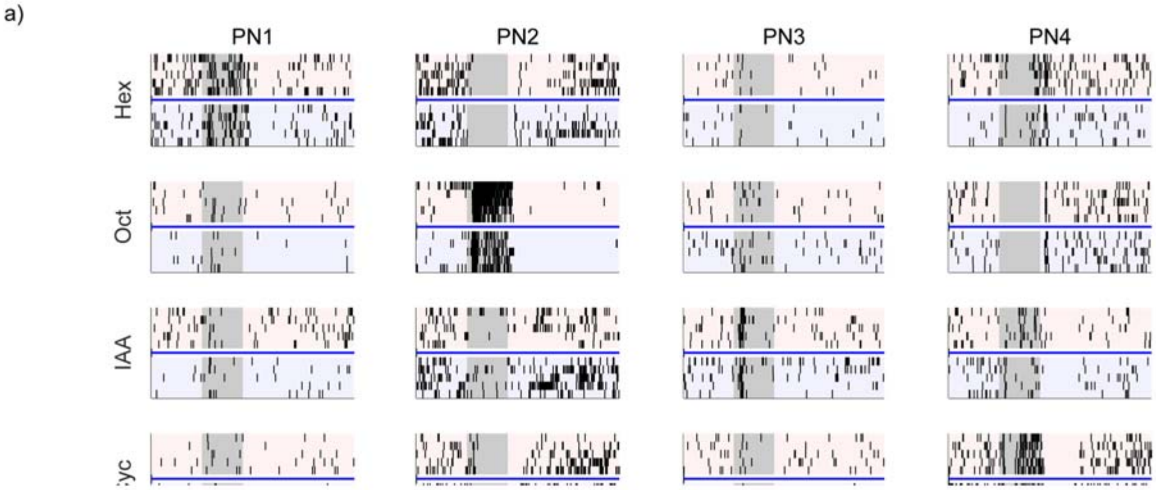
treatment, having severe or cerebral malaria, having received anti-malarial therapy in the previous week, known to have diabetes, chronic kidney disease or liver disease, or being uncooperative with breath collection were excluded from the study. Details about the children's age, gender, symptoms, and other factors are displayed in Figure 5.1. In brief, the average ages of children with malaria and without malaria were 8 and 7 respectively. Out of 35 total children, 31 had fevers, 30 had headaches, and 30 had abdominal pain.

Breath samples were analyzed with a GC/MS, after which traces were normalized to internal standards and contaminants were removed (See methods 2.4.2 and Figure 2.1.4). For visualization and examination of the collected breath profiles, the abundance of VOCs was plotted as a heatmap (Figure 5.2). The VOCs are plotted along the x-axis in order of elution time. Each row represents a different patient. Patients 1-18 are malaria negative while patients 19-35 are malaria positive. The colors represent the relative abundance for each VOC (note that blue indicates that the VOC had a very low abundance or was not present). The heatmap appears to show no clear differences in VOC composition across the two patient populations.

	Malaria Positive (*n* = 17)	**Malaria Negative (*n* = 18)**	***p* value**
Demographics			
Age, median years (IQR)	8 (6-10)	7 (5-8.5)	0.33
Female, *n* (%)	8 (47)	10/17 (59)	0.73
Reported Symptoms, *n* (%)			
Fever	16 (94%)	15 (83)	0.60
Diarrhea	0 (0%)	2 (11)	0.49
Vomiting	5 (29%)	4 (22)	0.71
Headache	16 (94%)	14 (78)	0.34
Abdominal Pain	13 (76%)	17 (94)	0.18
Muscle/Joint Pain	12 (71)	4 (22)	0.007
Other, *n* (%)			
Chronic Malnutrition*	5/16 (31)	3 (17)	0.43
Acute Malnutrition*	0/16 (0)	1 (6)	1
Use Bednet	9 (53)	10 (56)	1
Malaria within past 3 months	3 (18)	5/17 (29)	0.69

Figure 5.1. Patient population characteristics. Patient data are represented as number (%) except for age. IQR = interquartile range. If one or more patients were excluded, the number given is a fraction of total. Chronic and acute malnutrition defined respectively as height-for-age Z-score or BMI-for-age Z-score two or more standard deviations below median. Fisher's exact test or the Mann-Whitney U-test was used as appropriate to calculate p values. Table and legend are adapted from [5].

To better assess which biomarkers were indicative of malaria infection we computed correlations between the VOC profiles across patients and their known infection status, 1 for malaria (+) patients and 2 for malaria (-) patients (i.e. a class label for each breath sample). Correlations between the abundance of a VOC across samples with the assigned class labels ranged from -0.361 to 0.373. In a heatmap sorted by correlation values (Figure 5.3), we found that there was a distinction in VOC abundances between the two patient classes. The first 25 VOCs, those that are most positively correlated with class labels, have lower abundances for malaria (-) patients than malaria (+) patients. The reverse is true for the last 25 VOCs, those that are the most negatively correlated with the class labels.



While no individual VOC is indicative of malarial infection, and the same set of VOCs were not clearly present in all breath samples, we nevertheless hypothesized that the abundance profile across several VOCs could accurately predict malarial infection.

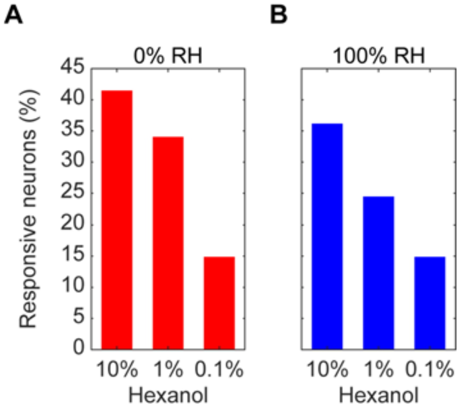


Figure 5.3. Sorted heatmaps of VOC abundance in patient breath. VOCs are sorted in order of decreasing correlation with patient labels. Y-axis indicates patient. Colors indicate abundance. Black line segments dividing malaria (-) patients (top) from malaria (+) patients (bottom). Subplots zoom into regions of high magnitude correlation. Cyan bordered box (left) shows positively correlated VOC where abundances are higher in malaria(+) patients. The red bordered box (right) shows negatively correlated VOCS where abundances are higher in malaria(-) patients.

To test this hypothesis we created a classifier based on cumulative abundance of the VOCs selecting those with highest correlation values with the class labels. The cumulative abundance was calculated by adding the abundance of positively correlated features and subtracting the abundance of negatively correlated features. Refer to Figure 5.4 for a schematic diagram of the data pre-processing and the cumulative abundance classifier. Classification was done using a nearest-mean classification approach (malaria (+) or malaria (-)).

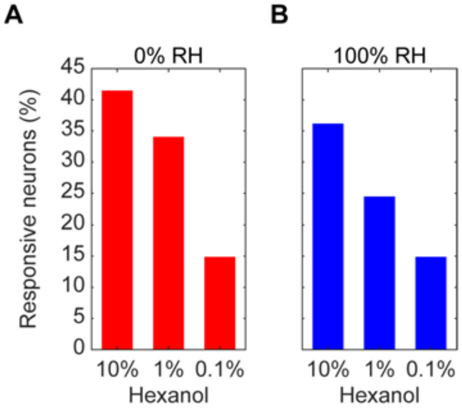


Figure 5.4 Schematic of data processing and analysis. A) Pre-processing steps for GC-MS spectra. B) Cumulative abundance classifier. VOCs with highest correlation magnitudes are selected for use in the classifier. Those with positive correlations are added and those with negative correlations are subtracted. Based on the final summation relative to a threshold, the patient is classified as malaria positive or malaria negative.

When using all the patients to calculate VOC correlations and pick biomarkers, we were able to achieve a classification rate of up to 94% incorporating just 25 VOCs with the highest absolute correlation values. This classification rate was relatively stable, because increasing the number of VOCs used, up to 100, did not change the classification rate significantly (Figure 5.5A). Since highly correlated features were selected first, each additional feature is less correlated than the last until we reach generally uncorrelated features. The abundances of these later features added no unique information to the classifier and did not substantially change the output classification; sometimes they even lowered it. We also found that, with the exception of 2 persistently falsely classified patients, correct classifications were common (Figure, 5.BB, patients 4 and 22). These results indicate that there is information that can be used to generally classify whether or not a patient is infected by malaria, based on the cumulative abundances of certain VOCs in exhaled breath.

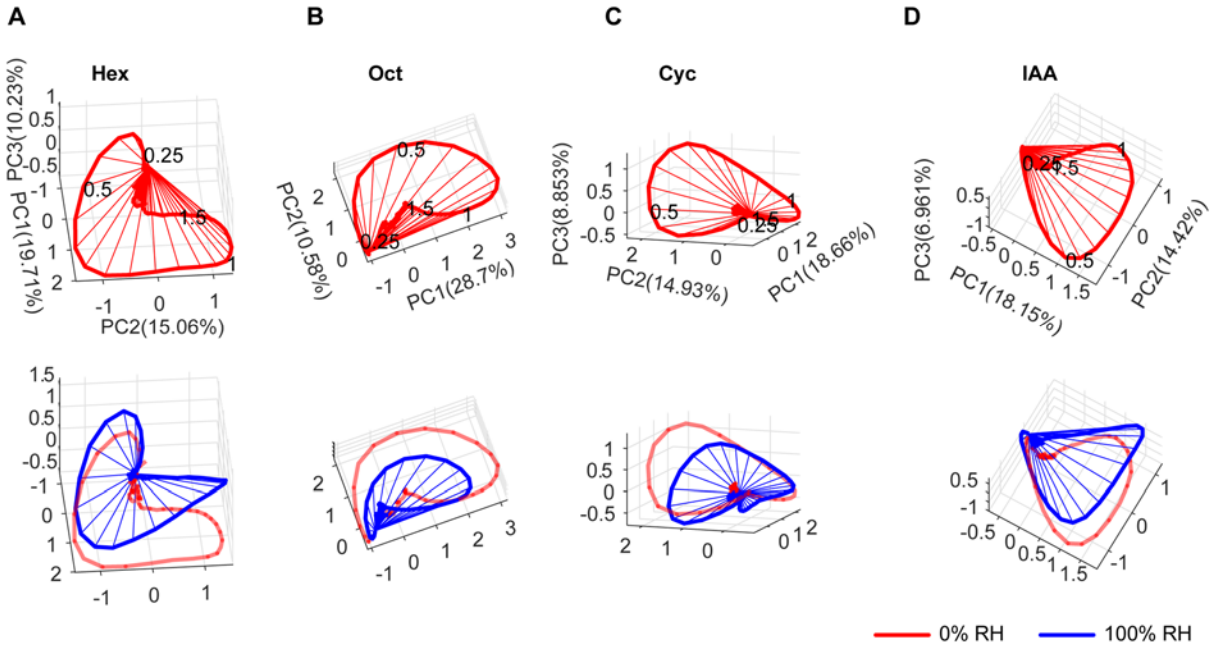


Figure 5.5. The cumulative abundance classifier yields highly accurate patient classifications. A) Classification rate as a function of number of markers used. Red point indicates first instance of peak classification. B) Classified malarial state for each patient as a function of

5.1.3 Validation of the cumulative abundance metric

A leave-one-patient-out validation was performed to assess the robustness of our classifier. We selected which VOCs to use in the classifier by using 34 out of 35 patients and used the resulting classifier to classify the single remaining patient. This process was repeated for all 35 patients. As can be expected, the classification performance reduced compared to the case when all the samples were used for feature selection and classification (Figure 5.5 C). Peak classification performance was only at 69%, much lower than the 94% seen earlier. Examining the change in features selected, we found that some features with relatively high correlation magnitudes were sparsely present in just a few patients (refer Figure 5.2-5.3). While this did not pose a problem when using all patients to generate a classifier in a leave-one-out or leave-k-out validation, the absence of sparse, yet highly correlated features, lowered classification accuracy.

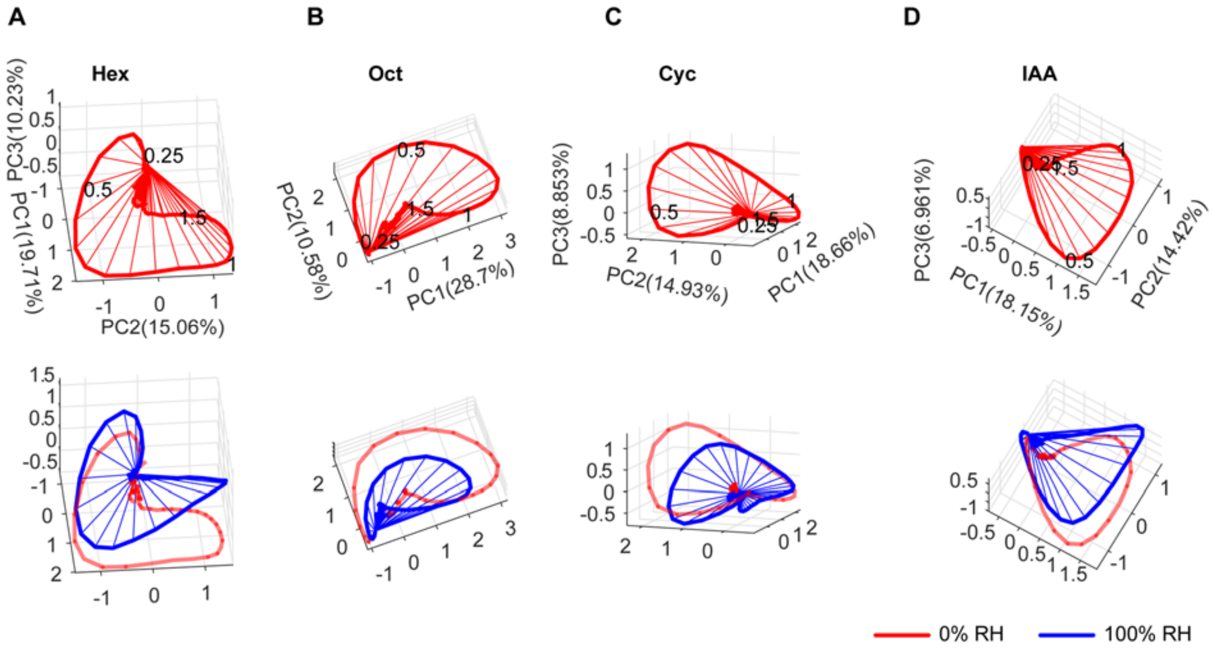
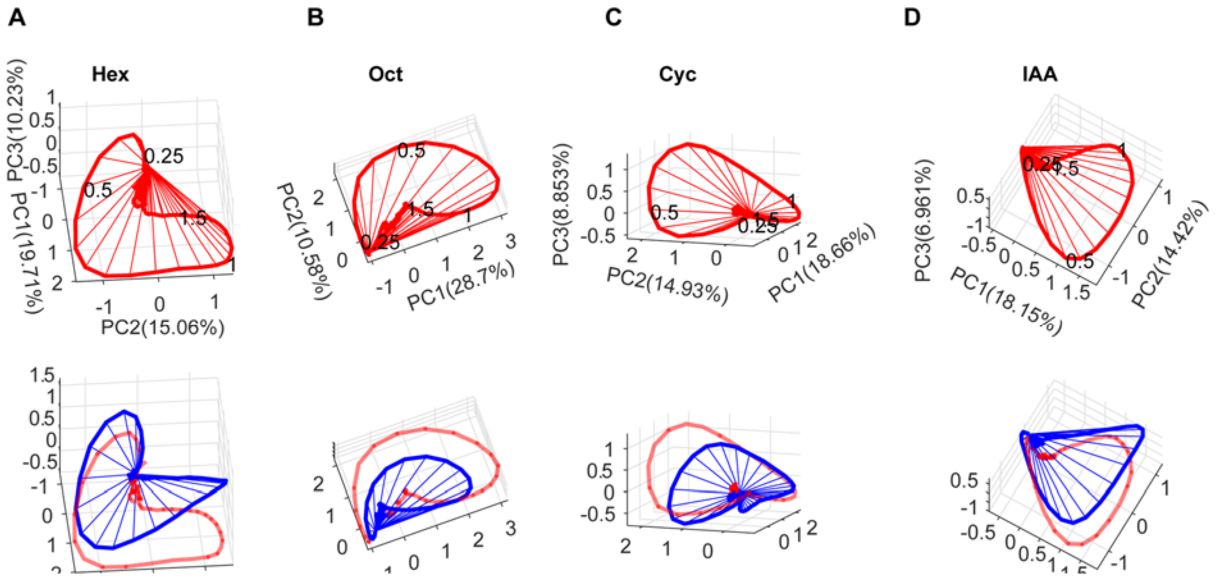


Figure 5.6. Eliminating sparse features increased classification validation performance. A)

To address the effects of VOC sparsity on our classifier, we explored the use of a threshold for the minimum number of patients in which a VOC had to be present to be considered a biomarker. The profile of features remaining as a function of the sparsity threshold is shown in Figure 5.6A. The number of VOCs decreased monotonically and reached a minimum of 7 features which were present in all 35 patients. A sparsity threshold of 20 was selected based on performance in the overall classification as well as the leave-out-validation (Figure 5.6 B and C). To be conservative, this point was local minimum in both plots and represented a number just over half the total patients. Note that correlations at this threshold ranged from -0.361 to 0.360, similar to the correlation range without a threshold.

We found that using a sparsity threshold to pre-screen features greatly improved classification performance. Overall classification reached a rate of 83%, and the leave-one-out validation rate of 77%, both using as few as 6 biomarkers. We found that 83% was a relatively stable classification rate with these parameters and that classification rates did not change significantly by adding more VOCs to the classifier (Figure 5.7 A). This stability is also apparent from the classification state heatmap (Figure 5.7 B) which shows that few patients change with an increasing number of VOCs.



Having selected optimal parameters for this classifier, we examined the abundances of the six most correlated biomarkers and how they varied with disease state. Whether or not a patient has malaria is marked by abundance decreases in exhaled breath of certain compounds (methyl undecane, nonanal, isoprene, tridecane) as well as increases of other compounds (dimethyl decane and trimethyl hexane, Figure 5.8A). Three of these compounds (marked by *) were identified through comparison with pure standards. The other three were manually annotated by reference to a spectral library. All of these compounds are known components of human breath [11]. Isoprene is endogenous in humans, and other VOCs are generated by oxidative stress-induced peroxidation [12, 13].

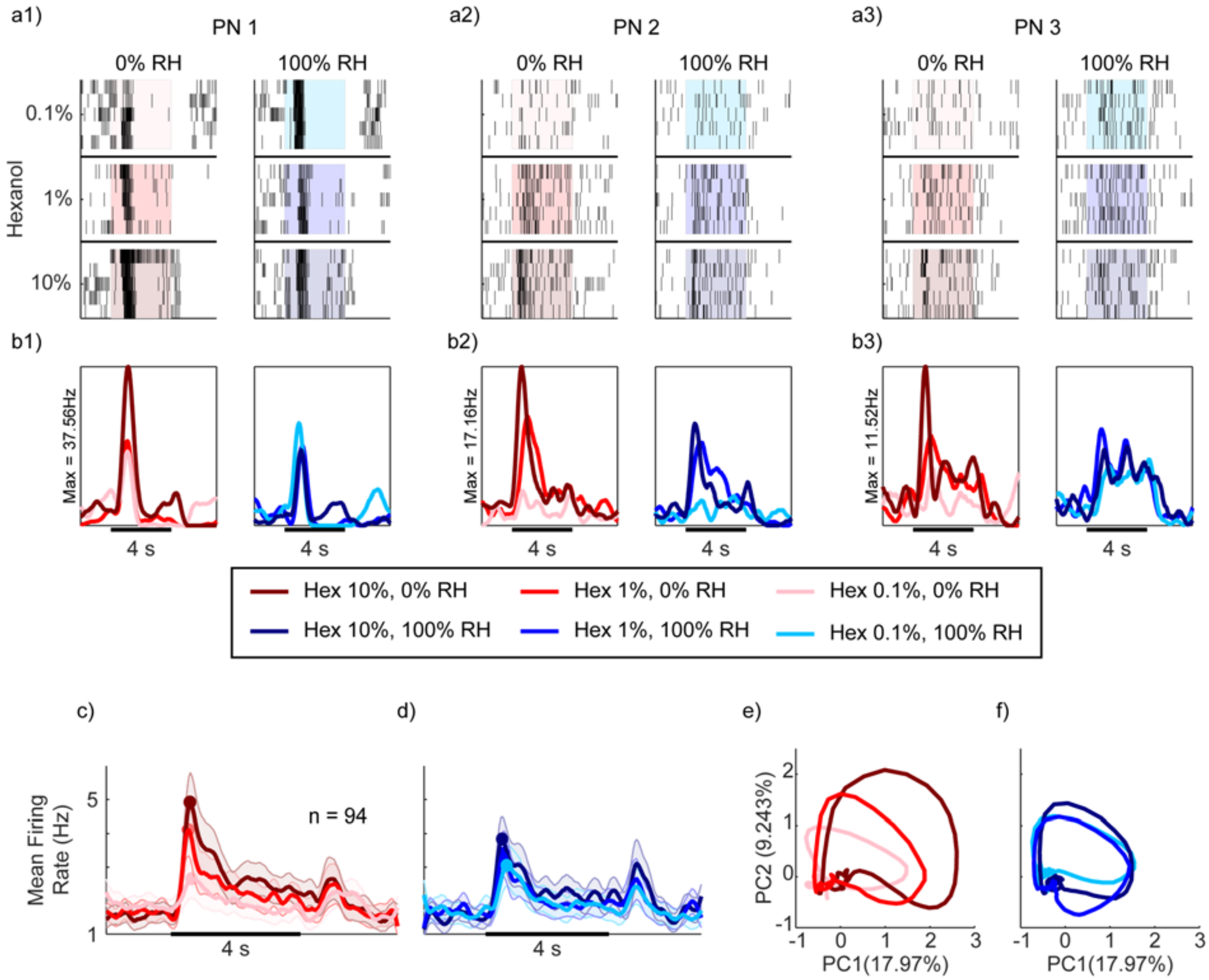


Figure 5.8. Classification of malarial state using VOCs abundant in at least 20 of 35 patients. A) Expression patterns of human breath biomarkers with largest absolute correlations to malarial state. * marks biomarkers identified by comparison with a pure standard. B)

The cumulative abundance metric provides a more nearly Gaussian distribution than individual component features, enabling better classification (Figure 5.8 B). Critically, with an appropriate cumulative abundance threshold, we found that as few as six VOCs correctly classified malaria infection status with 83% accuracy (Figure 5.8 C and D). Five patients who were actually malaria-positive were classified as being negative, and one patient who was malaria-negative was classified as being malaria-positive. Thus, the classifier has a specificity of 94% and sensitivity of 71%.

5.1.4 Diagnosing malaria

In previous studies, the presence of malaria has shown to increase the attractiveness of host vectors such as mosquitoes [14-18]. Specifically, 10-carbon monoterpenes, such as α-pinene, have been identified as influencing the attraction and feeding behavior of mosquitoes [17]. Our work complements these studies and identifies other compounds of interest as strong candidate markers for the diagnosis of malaria. All of these chemicals can potentially serve as biomarkers for diagnosis of malaria based on exhaled breath analysis. However, they impose different design constraints on the detection system. Given that malaria is endemic to resource poor regions with limited healthcare infrastructure, a diagnostic test would need to be inexpensive as well as simple to use and maintain. Our future work will focus on developing such sensors and their associated technologies to detect and identify malaria biomarkers at medically relevant concentrations.

5.2 References

1. WHO, *World Malaria Report 2016*, 2016, World Health Organization: Geneva. p. 186.

2. WHO. *Malaria*. 2016 December 2016 [cited 2017 April 17]; Available from: http://www.who.int/mediacentre/factsheets/fs094/en/.

3. Crawley, J., et al., *Malaria in children.* The Lancet, 2010. **375**(9724): p. 1468-1481.

4. Rieckmann, K.H., D.R. Davis, and D.C. Hutton, *PLASMODIUM VIVAX RESISTANCE TO CHLOROQUINE?* The Lancet, 1989. **334**(8673): p. 1183-1184.

5. Price, R.N., N.M. Douglas, and N.M. Anstey, *New developments in Plasmodium vivax malaria: severe disease and the rise of chloroquine resistance.* Current Opinion in Infectious Diseases, 2009. **22**(5): p. 430-435.

6. Miller, L.H., et al., *The pathogenic basis of malaria.* Nature, 2002. **415**(6872): p. 673-679.

7. Ashley , E.A., et al., *Spread of Artemisinin Resistance in Plasmodium falciparum Malaria.* New England Journal of Medicine, 2014. **371**(5): p. 411-423.

8. Mathison, B.A. and B.S. Pritt, *Update on Malaria Diagnostics and Test Utilization.* Journal of Clinical Microbiology, 2017.

9. WHO, *Guidelines for the treatment of malaria. Third Edition.* Geneva: World Health Organization, 2015: p. 316.

10. Schaber, C., et al., *Breathpriting Reveals Malaria-Associated Biomarkers and Mosquito Attractants.* Manuscript in preparation, 2017.

11. Costello, B.d.L., et al., *A review of the volatiles from the healthy human body.* Journal of Breath Research, 2014. **8**(1): p. 014001.

12. Hakim, M., et al., *Volatile Organic Compounds of Lung Cancer and Possible Biochemical Pathways.* Chemical Reviews, 2012. **112**(11): p. 5949-5966.

13. Mochalski, P., et al., *Blood and breath levels of selected volatile organic compounds in healthy volunteers.* Analyst, 2013. **138**(7): p. 2134-2145.

14. Lacroix, R., et al., *Malaria Infection Increases Attractiveness of Humans to Mosquitoes.* PLOS Biology, 2005. **3**(9): p. e298.

15. De Moraes, C.M., et al., *Malaria-induced changes in host odors enhance mosquito attraction.* Proceedings of the National Academy of Sciences, 2014. **111**(30): p. 11079-11084.

16. Berna, A.Z., et al., *Analysis of Breath Specimens for Biomarkers of Plasmodium falciparum Infection.* The Journal of Infectious Diseases, 2015. **212**(7): p. 1120-1128.

17. Kelly, M., et al., *Malaria Parasites Produce Volatile Mosquito Attractants.* mBio, 2015. **6**(2).

18. Emami, S.N., et al., *A key malaria metabolite modulates vector blood seeking, feeding, and susceptibility to infection.* Science, 2017. **355**(6329): p. 1076-1080.

Chapter 6: Conclusions and future work

Although there are many challenges to sensory detection in the chemical space, biological systems have evolved an exquisite system capable of overcoming them. In this work we took inspiration from biological olfactory systems to address three main challenges: overcoming sensor degradation, providing robust performance in the face of environmental variations, and detecting and identifying individual compounds.

6.1 Overcoming sensor drift

Sensor degradation is overcome in biological systems through constant regeneration of ORNs and their ability to re-innervate the same central circuitry. The turnover of olfactory receptor neurons has been well studied[1,] and has been shown across phyla in both simple organisms such as insects[2] and higher organisms such as mammals [1, 3, 4]. Remarkably, regenerated ORNs re-innervate the brain during regular regeneration of the olfactory sensors and primary signal processing cells as well as after injury [1-3, 5-7] or apoptosis. This is an attribute not found in any other sensory system. An important feature of both these features is that the central circuitry in the brain remains unaffected by changes at the sensory level, allowing for long term reliability of the system.

A number of sensing technologies have been proposed for non-invasive chemical sensing, yet they all share the fundamental problem of determining what features of the transducer's signal constitute a chemical fingerprint, allowing for precise analyte recognition. It is particularly important to be able to extract features that are robust with respect to the sensor's age or use, factors which cause drift in the signal. In this work we show that using pulsed stimulus delivery allows a sensor's operation to be modeled as a linear input/output (I/O) transform. The I/O transform is analyte specific and can be used to precisely predict a chemiresistor's response to the analyte, given the recent stimulus history. We have shown that the I/O transforms improve invariance to stimulus intensity and remain consistent despite considerable aging of the sensor. Significantly, the I/O transforms for a given analyte are highly conserved across sensors of equal manufacture, thereby allowing training data obtained from one sensor to be used for recognition of the same set of chemical species with another sensor. Thus, as do biological systems, this proposed approach facilitates decoupling of the signal processing algorithms from the chemical transducer, a key advance necessary for achieving long-term, non-invasive chemical sensing.

6.2 Invariance to environmental variation

In chemical sensing, especially vapor phase sensing, the environment is constantly fluctuating: The background gas will have varying components and concentrations, air flows will change in speed and direction, temperatures will fluctuate, and humidity levels will change. As seen in this work and published reports [8], humidity, in particular, has drastic effects on a chemical sensor signal. Nonetheless, biological systems are able to remain invariant to these fluctuations and robustly detect and recognize targets. In this work, we elucidated the central processing that enables this invariance, with the ultimate goal of replicating this technique in artificial olfactory systems.

We examined the effect of humidity on odorant identity and intensity coding. Our results reveal that, at the sensory level, humidity primarily affected the spontaneous baseline activity and had limited effects on stimulus intensity. The central circuits, on the other hand, had the reverse effects. There were limited changes to the spontaneous baseline activity, but stimulus intensity information was variably compressed in the antennal lobe. Our results also indicate that the odorant identity was robustly encoded, irrespective of the relative humidity level. Interestingly, the changes in the dynamic range of PN responses carried forward to behavior suggesting incomplete invariance of biological systems to humidity changes.

6.3 Specific detection

Biological systems rely on large arrays of odor-specific sensors to detect and discriminate among odorants. The locust antenna, for example, has approximately 50,000 individual sensors (ORNs) [2, 9]. Redundancy in this array improves the signal to noise ratio and facilitates detection over a large dynamic range. Many techniques have been employed in chemical sensing implementations with varying success, but as of yet nothing has come close to the versatility of biological olfactory systems in correctly identifying a large variety of odorants.

Biological sensors have also evolved to specifically sense biologically relevant compounds. In this work, we first developed a method to identify molecules relevant to the task of diagnosing malaria. Our cumulative abundance algorithm uses a GC/MS dataset to identify compounds indicative of a malaria infection. It incorporates the relative changes in VOC abundance in exhaled breath to make a diagnosis. We identified six compounds with different abundances in patients that were infected with malaria and those who were not.

While in this work we targeted malaria, which has large impacts on global health, this algorithm can be extended to any illness that would result in changes in exhaled breath composition due to metabolic health. Biomarkers have already been identified for several ailments, such as lung cancer [10] and fungal infections [11]. Having an expanded library of biomarkers and their abundance changes would enable the design of sensors targeting disease-specific biomarkers. The ultimate goal is the development of non-invasive diagnostic tools for a broad spectrum of human diseases and ailments.

6.4 Ongoing and future work

6.4.1 Detection of malaria biomarkers

The overarching goal of artificial olfaction has been the creation of a sensor that can detect and identify a wide range of chemical compounds with high specificity and sensitivity. Many groups have mimicked the size of biological sensor arrays to achieve this goal [12-18]. Here, instead, we explore an alternate approach of making a sensor specific to a narrowly defined problem, detecting biomarkers indicative of malaria. We show the use of a novel graphene-based chemiresistive sensor capable of specifically detecting low levels of a malaria biomarker. Targeting specific compounds in this manner allows for refinement of the sensor and opens the possibility of creating an array of several of such sensors for cross-selective sensing.

Studies have found that certain compounds, such as terpenes, are attractants to *Anopheles* mosquitoes, a primary vector for the spread of malaria [19, 20]. Other studies have shown that α-pinene is produced by *P. falciparum*-infected erythrocytes, and directly activates *Anopheles* odorant receptors (AgOR21 and AgOR50), and modulates mosquito feeding behavior [21, 22].

Malaria is endemic to and primarily found in Southeast Asia and sub-Saharan Africa. These regions have limited resources and infrastructure. Ultimately, a diagnostic sensor deployed there should have low manufacturing costs and be simple to manufacture and use. Typical metal-oxide sensors, such as the one discussed in Chapter 3, require built-in heaters which add complexity and expense to their design and use. Here we show the use of a novel graphene based chemiresistive sensor capable of detecting low levels of α-pinene without a heater.

Details of the manufacture of the sensing film and sensor can be found in the methods, but are briefly discussed here. We micro-fabricated an interdigitated electrode array. Crumpled graphene oxide was provided to us by collaborators[3] who followed previously published manufactured methods [23] and was provided to us by. The cGO was dropcast onto the sensor until a resistance of ~1 MΩ was achieved. The sensor was then wired and packaged on a DIP socket for use.

6.4.2 Sensor Responses to malaria α-pinene

We probed the response of this sensor with four analytes: 1-octanol, 2-octanol, geraniol, and α-pinene. We selected 1- and 2-octanol because of their structural similarity to see if the sensor could differentiate such structurally similar compounds. Similarly, geraniol and α-pinene were selected because they are both plant-derived terpenes. Representative responses to these compounds are shown in Figure 6.1. The sensor was exposed to successive stimulus pulses with each pulse greater in concentration than the last. Stimulus periods were 5 minutes in duration and there were 5 minutes between successive stimulations.

When the VOC stimulus is presented to the sensor (gray shading), the bandpass filtered (1-10 mHz) film resistance, represented by signal-to-noise ratio (SNR, R/R_{STD}), rapidly increases. The response quickly reaches as steady state, a level at which the activity is steady or very slowly changing. Once the stimulus is removed, the signal rapidly returns to or near baseline levels. We observed strong concentration-dependent behavior for three of the four analytes: geraniol, 1-octanol, and 2-octanol (Figure 6.1A-C). However, α-pinene had an inverse relationship with concentration. Each successive stimulation resulted in a weaker response magnitude (Figure 6.1D). The mechanism behind the response behavior to α-pinene is yet to be

---

[3] Crumpled graphene oxide was provided to us by our collaborator Dr. Ramesh Raliya in the group of Dr. Pratim Biswas. Details of its manufacture and morphology are published as: Wang, W.-N., Y. Jiang and P. Biswas (2012). "Evaporation-Induced Crumpling of Graphene Oxide Nanosheets in Aerosolized Droplets: Confinement Force Relationship." The Journal of Physical Chemistry Letters **3**(21): 3228-3233.

determined. Interestingly, despite the weakening response with each pulse, the signal consistently returned to baseline levels with time constants similar to those of responses to other concentrations and analytes, suggesting that it is not a simple case of saturation or poisoning.

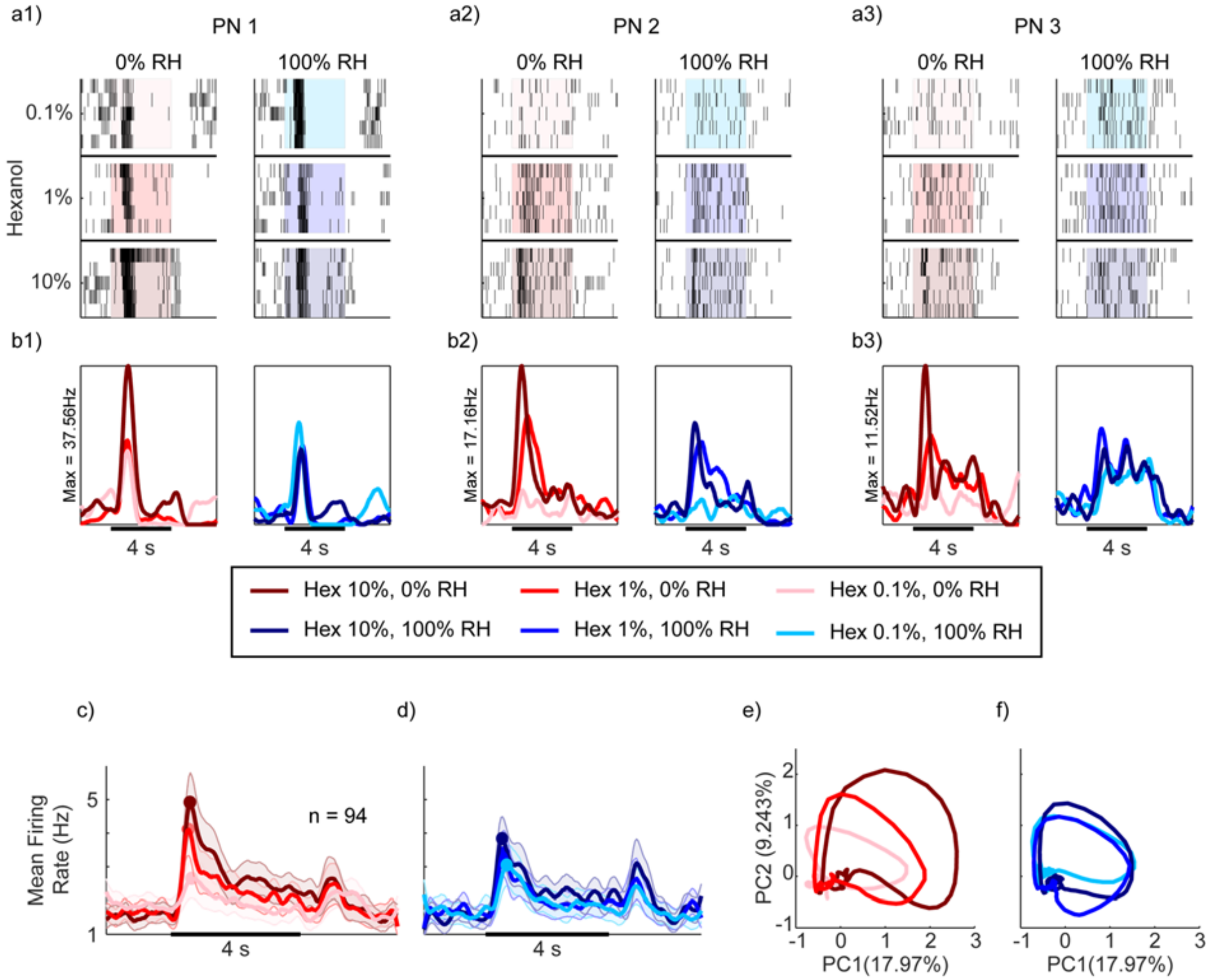
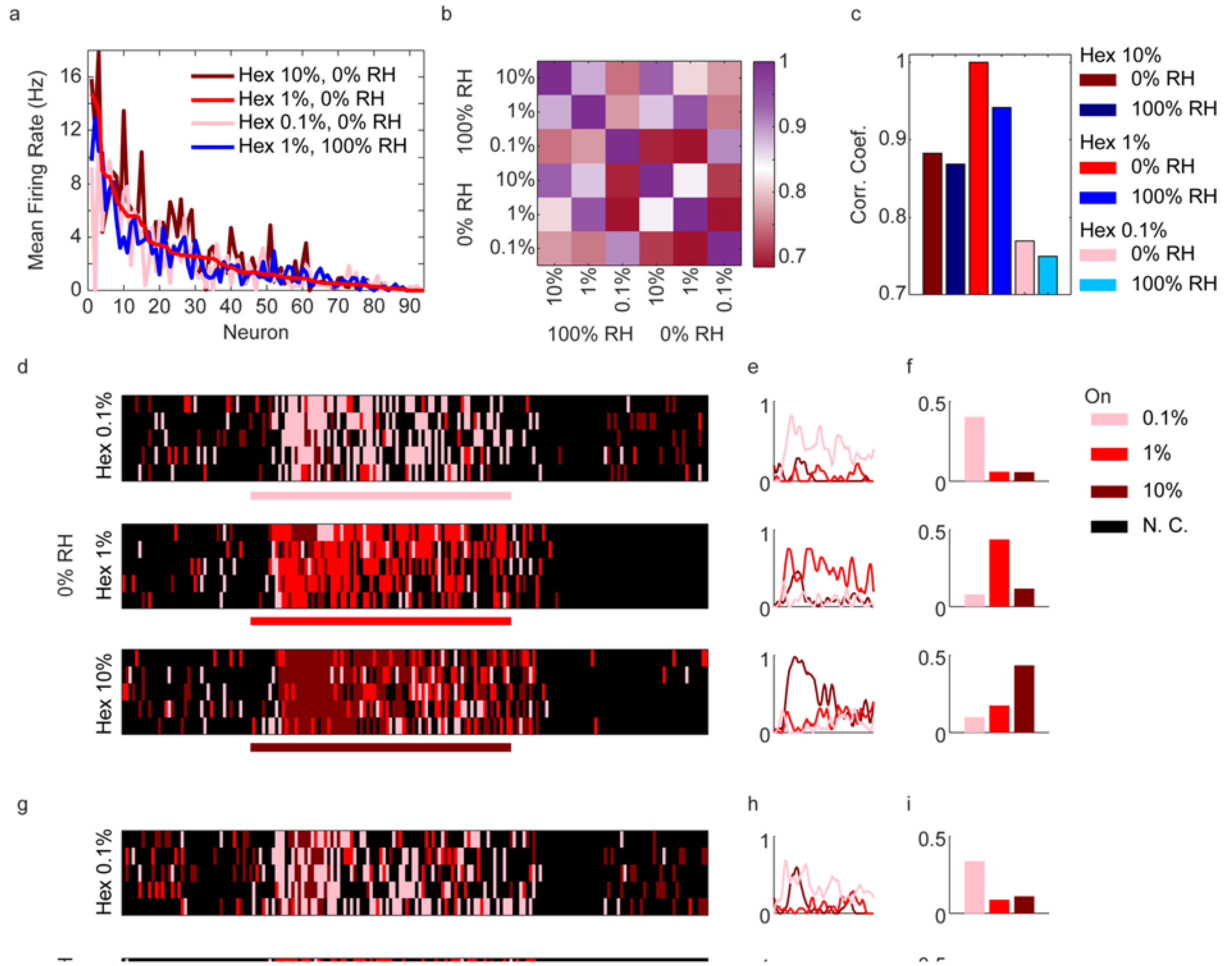


Figure 6.1. Representative responses of a cGO based sensor to four different analytes. Four pulses of each analyte are presented, with each pulse increasing in concentration. Gray bars indicate periods of analyte exposure. Concentrations for each analyte (in PPM) are as follows: geraniol: 0.02, 0.05, 0.07, 0.09; 1-octanol: 0.4, 0.8, 1.1, 1.5; 2-octanol: 0.8, 1.5, 2.2, 2.9; α-pinene: 15.5, 30.2, 44.3, 57.8. Signal from sensors was band-pass filtered between 1 and 10 mHz.

6.4.3 Limits of detection for a cGO based sensor

Interestingly, the sensitivity of this sensor is strongly analyte dependent. In assessing the limit of detection (LOD), we found that 1- and 2-octanol had similar sensitivities, with detection limits of 1.5E-1 PPM (Figure 6.2 A-B). Geraniol, however, evokes response magnitudes of similar strength (SNR of 10-30) to those of other analytes tested, but at concentrations 1-2 orders of magnitude lower, with a LOD of 8E-3PPM (Figure 6.2C). Conversely, we found that α-pinene evoked responses at concentrations 1-2 orders of magnitude higher than the other analytes, with a LOD of 7.7 PPM. The differences in sensitivity are clear from the LOD for each analyte. This material is most sensitive to geraniol and least to α-pinene.



6.4.4 Selective sensing of α-pinene

Designing a sensing system for biomarker detection requires both sensitivity and specificity. Although the cGO film was least sensitive to α-pinene, there was a unique saturation effect which could be used to impart high levels of specificity. The response magnitude of the sensor decreased to successive pulses of α-pinene. The exact reason for this effect is uncertain, but it opens the possibility of selectively sensing this analyte based on hysteresis. We hypothesized that repeated presentations of α-pinene would decrease response magnitudes until little to no response was elicited. Other analytes, however, would not be subject to this phenomenon and would have similar responses to repeated stimulations. To test this, we performed back-to-back trials of four pulses with increasing concentrations for each analyte (Figure 6.3). Overlaying these trials, we find that the responses are very consistent in shape and magnitude for geraniol, 1-octanol, and 2-octanol (Figure 6.3 A-C left panel). The responses to α-pinene, however, see a great decrease in strength over time, and there is almost no response in the second trial (Figure 6.3 D, left panel). Subtracting the repeated trials, we find that the resulting signal is canceled out for all the analytes except α-pinene (Figure 6.3, right panel). Thresholding the resulting response difference provides a selective indicator for the presence of α-pinene (Figure 5.11 D, right panel).

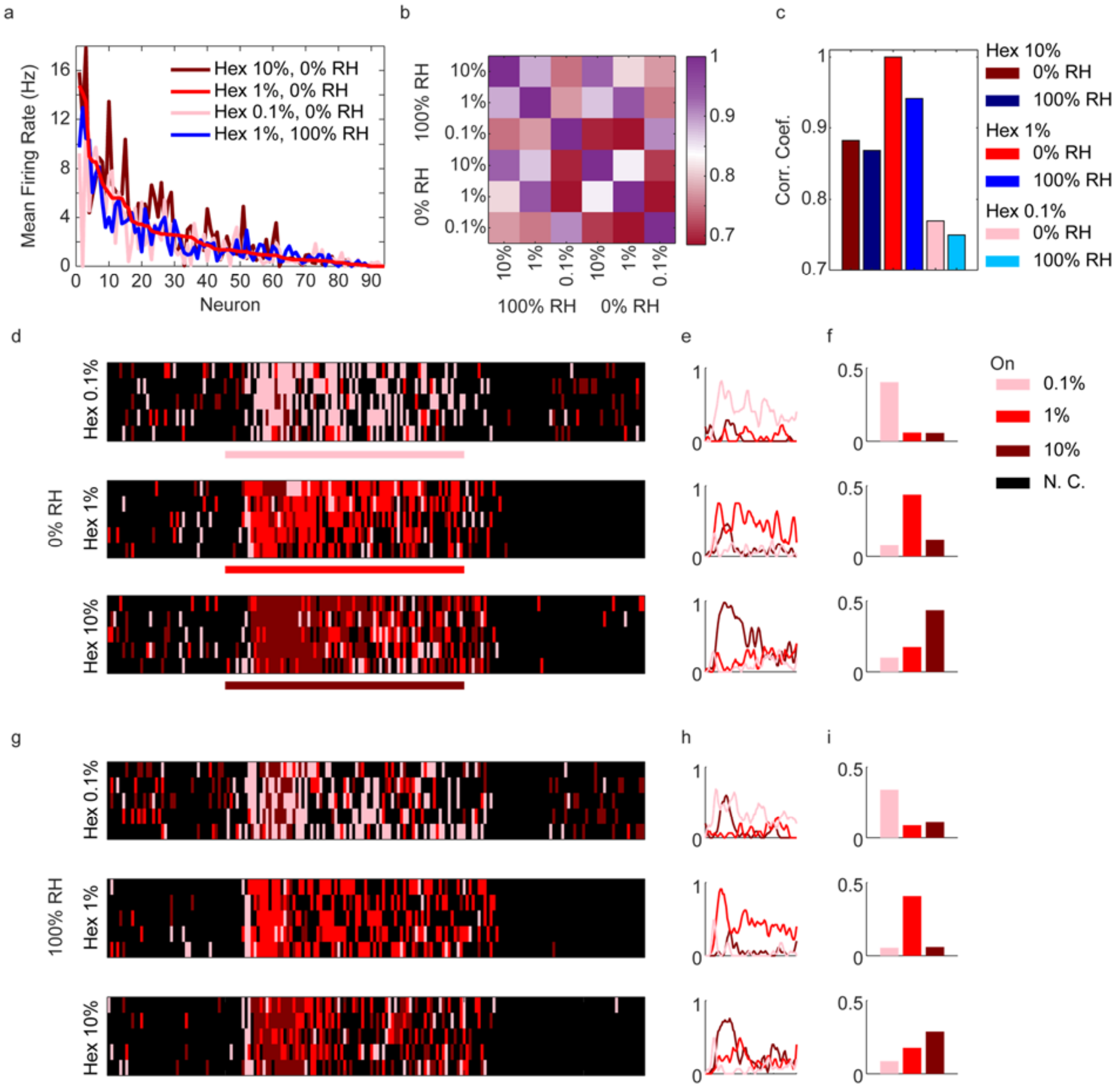


Figure 6.3. Sensor response for 1-octanol, 2-octanol and geraniol cancels out in back-to-back trials, but not for alpha-pinene. Left plots show two, back-to-back trials of analyte exposures. Right plots show the difference between the two trials. Gray bars reflect analyte exposure. Dotted red line is a threshold indicating when α-pinene is detected. Concentrations for each analyte (in PPM) are as follows: Geraniol: 0.008, 0.01, 0.02, 0.03; 1-octanol: 0.07,0. 2, 0.2, 0.3; 2-octanol: 0.4, 0.8, 1.1, 1.5; Pinene: 15.5, 30.2, 44.3, 57.8. Signal from sensors was band-

6.4.5 Limitations of cGO sensor for α-pinene detection

This cGO based sensor shows promising results for selectively sensing α-pinene. Nonetheless, limitations remain which must be addressed before a sensor based on this technology can be deployed for actual diagnostic purposes. There two most salient issues which concern the properties of exhaled breath. The first is that exhaled breath is highly humidified, with reported RH levels up to 80% [24]. Our sensor, however, fails under humidified conditions (Figure 6.4). Through testing, we found that the addition of humidified air increases film resistance and prevents the interaction of analytes that would introduce a response. The other issue is that our sensor is increasingly insensitive to α-pinene at biological relevant ranges. While the LOD for α-pinene established in this work is 7.7 PPM, actual concentrations calculated from the breath samples we collected were in the hundreds of parts-per-trillion (PPT) range. There are readily available solutions to address these issues. Breath samples could be pre-concentrated prior to analysis, which would remove humidity as a factor as well as address the sensitivity issue. This, however, would increase complexity and cost of the testing and thus eliminate some of the benefits of this test over existing methods.

Other limitations of this sensor that still need to be assessed are the reliability of response and its longevity. Reliability of response is particularly important in diagnostic applications. In our testing, we found some instances of unexplained variations in the response strength. These need to be explored in detail and resolved. We also need to verify the longevity of the sensor. It is important to know how long or how often the sensor can be used before response strength and reliability are affected. Given the relatively simplicity of this sensor and the film, we presume it will have minimal production costs and can be designed for single use testing.

[1]

Addressing the limitations of this sensor would open exciting possibilities for non-invasive diagnostics. For malaria in particular, there is a need to improve low-cost testing modalities because the parasite is primarily endemic to the developing world. Here we have presented an alternative indicator for the presence of malaria in pediatric patients, as well as an exciting new gas sensing film that is specific to this analyte.

6.5 References

1. Graziadei, P.P.C. and G.A.M. Graziadei, *Neurogenesis and neuron regeneration in the olfactory system of mammals. I. Morphological aspects of differentiation and structural organization of the olfactory sensory neurons.* Journal of Neurocytology, 1979. **8**(1): p. 1-18.

2. Stern, M., et al., *Regeneration of olfactory afferent axons in the locust brain.* The Journal of Comparative Neurology, 2012. **520**(4): p. 679-693.

3. Graziadei, P.P.C. and G.A.M. Graziadei, *Neurogenesis and Plasticity of the Olfactory Sensory Neuronsa.* Annals of the New York Academy of Sciences, 1985. **457**(1): p. 127-142.

4. Riecke, H., *Olfactory Computation and Adult Neurogenesis.* 2013.

5. Gheusi, G. and P.-M. Lledo, *Control of Early Events in Olfactory Processing by Adult Neurogenesis.* Chemical Senses, 2007. **32**(4): p. 397-409.

6. Hinds, J.W. and P.L. Hinds, *Synapse formation in the mouse olfactory bulb. II. Morphogenesis.* The Journal of Comparative Neurology, 1976. **169**(1): p. 41-61.

7. Hinds, J.W. and P.L. Hinds, *Synapse formation in the mouse olfactory bulb Quantitative studies.* The Journal of Comparative Neurology, 1976. **169**(1): p. 15-40.

8. Williams, D.E., *Semiconducting oxides as gas-sensitive resistors.* Sensors and Actuators B: Chemical, 1999. **57**(1–3): p. 1-16.

9. Laurent, G., *Dynamical representation of odors by oscillating and evolving neural assemblies.* Trends in Neurosciences, 1996. **19**(11): p. 489-496.

10. Hakim, M., et al., *Volatile Organic Compounds of Lung Cancer and Possible Biochemical Pathways.* Chemical Reviews, 2012. **112**(11): p. 5949-5966.

11. Koo, S., et al., *A Breath Fungal Secondary Metabolite Signature to Diagnose Invasive Aspergillosis.* Clinical Infectious Diseases, 2014. **59**(12): p. 1733-1740.

12. Cavicchi, R.E., et al., *Optimized temperature-pulse sequences for the enhancement of chemically specific response patterns from micro-hotplate gas sensors.* Sensors and Actuators B: Chemical, 1996. **33**(1–3): p. 142-146.

13. Kunt, T.A., et al., *Optimization of temperature programmed sensing for gas identification using micro-hotplate sensors.* Sensors and Actuators B: Chemical, 1998. **53**(1–2): p. 24-43.

14. Boger, Z., et al., *Rapid Identification of Chemical Warfare Agents by Artificial Neural Network Pruning of Temperature-Programmed Microsensor Databases.* Sensor Letters, 2003. **1**(1): p. 86-92.

15. Vergara, A., et al., *Optimized temperature modulation of micro-hotplate gas sensors through pseudorandom binary sequences.* Sensors Journal, IEEE, 2005. **5**(6): p. 1369-1378.

16. Benkstein, K., et al., *Integration of nanostructured materials with MEMS microhotplate platforms to enhance chemical sensor performance.* Journal of Nanoparticle Research, 2006. **8**(6): p. 809-822.

17. Vergara, A., et al., *Quantitative gas mixture analysis using temperature-modulated micro-hotplate gas sensors: Selection and validation of the optimal modulating frequencies.* Sensors and Actuators B: Chemical, 2007. **123**(2): p. 1002-1016.

18. Rogers, P.H. and S. Semancik, *Development of optimization procedures for application-specific chemical sensing.* Sensors and Actuators B: Chemical, 2012. **163**(1): p. 8-19.

19. Nyasembe, V.O., et al., *Behavioural response of the malaria vector Anopheles gambiae to host plant volatiles and synthetic blends.* Parasites & Vectors, 2012. **5**(1): p. 234.

20. Nikbakhtzadeh, M.R., et al., *Olfactory basis of floral preference of the malaria vector Anopheles gambiae (Diptera: Culicidae) among common African plants.* Journal of Vector Ecology, 2014. **39**(2): p. 372-383.

21. Kelly, M., et al., *Malaria Parasites Produce Volatile Mosquito Attractants.* mBio, 2015. **6**(2).

22. Emami, S.N., et al., *A key malaria metabolite modulates vector blood seeking, feeding, and susceptibility to infection.* Science, 2017. **355**(6329): p. 1076-1080.

23. Wang, W.-N., Y. Jiang, and P. Biswas, *Evaporation-Induced Crumpling of Graphene Oxide Nanosheets in Aerosolized Droplets: Confinement Force Relationship.* The Journal of Physical Chemistry Letters, 2012. **3**(21): p. 3228-3233.

24. Peng, G., et al., *Diagnosing lung cancer in exhaled breath using gold nanoparticles.* Nat Nano, 2009. **4**(10): p. 669-673.