Optical sensor arrays for chemical sensing: the optoelectronic nose

Jon R. Askim,*a Morteza Mahmoudiab and Kenneth S. Suslick*a

A comprehensive review is presented on the development and state of the art of colorimetric and fluorometric sensor arrays. Optical arrays based on chemoresponsive colorants (dyes and nanoporous pigments) probe the chemical reactivity of analytes, rather than their physical properties. This provides a high dimensionality to chemical sensing that permits high sensitivity (often down to ppb levels), impressive discrimination among very similar analytes and exquisite fingerprinting of extremely similar mixtures over a wide range of analyte types, both in the gas and liquid phases.

1 Introduction and classes of chemical sensors

As a species, we are visual creatures and underappreciate the olfactory sense. Nonetheless, even humans can recognize and discriminate more than 10000 different odorants. Molecular recognition by the primary olfactory system derives its specificity in an entirely different manner than the usual lock-and-key substrate–enzyme specificity. Instead, olfactory specificity originates from pattern recognition of the responses of several hundreds of highly cross-reactive olfactory receptors. Indeed, for land-based animals, there are typically about 1000 active olfactory receptor genes, which represents roughly 3% of our entire genome.2,3

Development of rapid, sensitive, portable and inexpensive systems for identification of a wide range of toxic gases, vapors, and aqueous solutions has become an urgent societal need and has important applications ranging from the chemical workplace to the general population. New approaches to chemical sensing4–13 with improved discriminatory powers are essential...

Jon R. Askim grew up in Vancouver, WA, and received his BS in Chemistry from Western Washington University. He is currently completing his PhD in the Department of Chemistry of the University of Illinois at Urbana-Champaign in the Suslick Research Group.

Morteza Mahmoudi received his PhD from Institute for Nanoscience and Nanotechnology at Sharif University of Technology (in collaboration with UC Davis, MIT, UBC, and EPFL) with specialization on the biomedical applications of superparamagnetic iron oxide nanoparticles. He is Director of NanoBio Interaction Laboratory at Tehran University of Medical Sciences (http://www.biospion.com). His current research involves control of protein corona decoration at the surface of nanoparticles and hidden parameters that affect the nanobio interfaces. He was a visiting professor at the University of Illinois at Urbana-Champaign working on detection of nanoparticles using optoelectronic sensor arrays.
to eliminate false positives for the monitoring of toxic gases at sub-ppm levels. In 1982, Persaud and Dodd tried to mimic the olfactory system using semiconductor transducers; this was one of the first artificial devices to successfully discriminate among a wide variety of odors without the use of highly specific receptors, i.e., an electronic nose.

The long history of visual indicators in analytical chemistry has led to the development of optical sensor arrays as an alternative to electronic sensors. There has been much recent progress in this “optoelectronic nose” approach, and we will present here an overview of the recent progress in colorimetric and fluorometric sensor arrays, examine the methods of analysis of the high dimensional data so obtained, and review important and diverse applications. We also discuss previous limitations of sensor arrays and prior electronic nose technology and comment on the recent successes in overcoming those failings. Finally, we will examine the emerging trends that are likely to impact the development of new optical sensor arrays.

Optical sensor arrays provide a facile, efficient, and sensitive approach for the rapid detection and identification of wide range of chemical substrates based on colorimetric or fluorescent changes quantified by digital imaging. Every optical sensor array must contain both an active center that can interact strongly with desired analytes and an intense chromophore or fluorophore that is strongly coupled to that active center. It is the intermolecular interactions of analytes with the active center, often through strong chemical interactions rather than simple physical adsorption, that results in a colorimetric or fluorometric change (i.e., chemoresponsive). Using a chemically diverse array of such chemo-responsive colorants, one generates a pattern that is an optical fingerprint for any odorant or mixture of odorants. The colorants are cross-reactive, but the pattern of the array response is unique: in this manner, olfactory-like responses are converted into an easily monitored optical output, thus acting as an optoelectronic nose.

1.1 Chemical sensors

There is an increasing demand to measure the chemical environment both inside our own bodies and in our surrounding environment. The healthy function of the human body is dependent on various chemical processes; thus, in situ monitoring is of crucial importance, e.g., the chemical composition of patient blood samples in medical diagnosis. The impact of the environment on living systems similarly is of critical interest, and analysis of types and amounts of a wide range of toxins, pollutants, and even naturally occurring chemical species is of increasing importance.

Such analyses and diagnostic methods need suitable sensors. A sensor is a device that converts an input signal into a readable output signal. The input signal can be any measurable characteristic such as quantity or physical variation, while the output is ultimately an electrical signal. Small and inexpensive sensors enable mass production and widespread application. Indeed, the development of new sensor technology faces the dilemma of trying to create sensors that are both increasingly sensitive and increasingly robust. Just as position and momentum are canonical variables, one may argue that beyond a certain point, the more sensitive a sensor becomes, inherently the less robust it can be. As we shall see, one path around this dilemma is the development of disposable sensors, thus unlinking the opposing demands.

Chemical sensors respond to the chemical environment (i.e., interactions with molecular species), rather than the physical environment (e.g., temperature or pressure). Chemical sensors can therefore be categorized into two major groups: those that discriminate among analytes based on physical properties (e.g., molecular weight, vapor pressure, etc.) and those that measure chemical properties (e.g., reactivity, redox potential, acid–base interactions, etc.). Chemical sensors can also be grouped by their signal transduction methods into three classes: (1) electrical and electrochemical, (2) thermometric, and (3) optical. We will focus on this last class of sensors as array components, but let us first briefly overview all three sensor transduction classes.

1.2 Electrical and electrochemical sensors

An electrical sensor is a resistive or capacitive measurement device that responds to analyte interactions with receptor layer of sensor. In a sense, the olfactory receptors of the vertebrate olfactory system are a large array of bioelectrical sensors. In chemical sensing, electrical and electrochemical sensors are nearly indistinguishable. Both intrinsically involve the interaction of chemical analytes with an electrical circuit and resistance, capacitance, current, or voltage as the monitored response. In attempts to mimic the olfactory system, a wide variety of electrical and electrochemical sensors have been explored, including metal oxide semiconductors (e.g., SnO₂), metal oxide semiconductor field effect transistors (FET) [e.g., Ga₂O₃] and chemical field-effect transistors (ChemFET) of both intrinsically conductive (e.g., polythiophenes) and

Kenneth S. Suslick

Kenneth S. Suslick received his BS from Caltech in 1974, his PhD from Stanford in 1978, and came to the University of Illinois at Urbana-Champaign immediately thereafter, where he is presently the Marvin T. Schmidt Professor of Chemistry. He is a fellow of the ACS, MRS, AAAS, and ASA. Among his awards are the ACS Nobel Laureate Signature and Senior Cope Scholar Awards, the ISOEN Wolfgang Göpel Award, and the RSC Sir George Stokes Medal. His two major research areas are the chemical and physical effects of ultrasound and chemical sensing and olfaction, both biological and artificial.
composite types (e.g. carbon-black polymer composite), 12 and coated quartz vibrators or acoustic wave sensors.42 In general, electrical sensors rely fundamentally on physical adsorption as the primary interaction between the chemical analytes and the electrically active surface; this reliance on physical adsorption, however, inherently provides for sensitivity to change in humidity, which remains a huge problem for both laboratory and especially field use. 6,8,9,13,28 In addition, aging of the surface of electrical sensors can induce significant baseline drift.6 Most recently, tremendous efforts have been developed with low dimensional nanomaterials as highly sensitive sensor transducers: e.g., nanowires, nanotubes, nanofibers, graphene and single (or few) layer two-dimensional materials.43–51 In large part, the potential for high sensitivity comes from the extreme surface area to mass ratios intrinsic to low dimensional materials, which permit extraordinary sensor exposure to analyte interactions. Selectivity remains key, however, to the future success of such systems for chemical or biomedical sensor applications.

Electrochemical sensors 52–55 are also extremely diverse and can be used for a wide range of applications, but are beyond the scope of this review. The reader is referred to recent reviews on applications in environmental, 56–60 clinical, 61–63 biosensing, 39,55,64–67 food 68 and biohazard agent 69 analyses.

1.3 Thermometric sensors

The mechanism of transduction in thermometric sensors relies on the measurement of the local heat change from specific chemical reactions or adsorption events involving the analyte. Thermometric sensors are constructed by coating a catalytic sensor layer on the surface of a thermometer. In this case, the interaction of a target analyte with a chemical sensor can generate or consume heat that is then measured by sensitive thermistor (i.e., semiconductors with strongly temperature-dependent conductivity).70,71 Thermometric sensors most commonly utilize enzymatic reactions with high enthalpy changes.28,72–74 Due to the simplicity of the thermal biosensing approach (e.g., there is no need for labeling reactants), this method can be considered as a suitable replacement for other signal transduction methods that require a sophisticated cascades of reaction steps.

A wide range of applications (e.g., detection of sucrose,75 glucose,76–79 uric acid, 80 insulin, 81 and lactate 82 by suitable immobilized enzymes) have been reported for thermometric biosensors. Moreover, a multi-analyte determination method has been performed by thermal biosensors using MEMS thermopiles. 78,83,84

1.4 Optical sensors

Optical sensors use visible or ultraviolet light to interrogate sensors for analysis. Optical sensors can be represented in general terms as a wavelength selectable light source, the sensor material itself interacting with analytes, and a light detector (Fig. 1). What the detector monitors varies by technique (e.g. refractive index, scattering, diffraction, absorbance, reflectance, photoluminescence, chemiluminescence, etc.), can cover different regions of the electromagnetic spectrum, and can allow measurement of multiple properties (e.g. intensity of light, lifetime, polarization, etc.).28,85 The focus of this review, however, is on optical sensor arrays that use absorbance, reflectance or fluorescence array detectors (i.e., digital cameras or scanners).

Colorimetry (i.e., quantitative measurement of absorbance or reflectance spectra) is, of course, one of the oldest of analytical techniques,46 and colorimetric sensors stretch back even before the beginnings of chemistry (e.g., squeeze a lemon into tea) with straightforward “naked-eye” quantitation. Colorimetric detection is a fairly simple technique, and the advent of universal digital imaging has given it new and exciting possibilities. We will use the general term colorimetry to include simple three color (i.e., RGB) imaging, hyperspectral imaging (i.e., more color channels), and full spectrophotometry (i.e., hundreds of color channels with nm resolution).

Fluorometry (i.e., the quantitative measurement of fluorescence spectra, cf. Fig. 2) 87–89 can provide excellent sensitivity and fluorescent sensors often have some advantages (e.g., sensitivity, depending upon the background fluorescence), although at the cost of a more sophisticated experimental apparatus.90 Fluorescence-based approaches and fluorescence parameters (e.g. Stokes shift, fluorescence intensity and anisotropy, emission and excitation spectra, and fluorescence lifetime) can provide substantial flexibility as an analytical approach.91,92 Fluorescence techniques can be divided into three main classes: intrinsic fluorescent,93–96 extrinsic fluorescent,91,97,98 and displacement or differential99–103 probes.

Fig. 1 General arrangement of spectroscopic measurements: (A) light reflection; (B) light refraction; (C) light absorption; (D) fluorescent emission. Reproduced with permission from ref. 85.

Fig. 2 Partial Jablonski diagram for absorption, fluorescence, and phosphorescence. Reproduced with permission from ref. 104.
Optical chemical sensors must perform two functions: they must both interact with analytes and subsequently report on such interactions (e.g., by changing color). While many dyes and fluorophores do so intrinsically, there are also many other “artificial receptors” (i.e., compounds capable of supramolecular interactions) that are not spectroscopically active. For example, crown ethers, cryptands, cyclodextrins and calixarenes often have excellent molecular recognition capabilities to selectively bind analytes of interest (especially cations, often anions, and sometimes neutral organics), but they are often spectroscopically inert. Such complexing agents can be covalently modified to incorporate a suitable chromophore or fluorophore, which can then report on analyte binding.

For example, the linkage of an anthracene fluorophore with a crown ether receptor creates a diamine sensor for detecting food spoilage. In the absence of guest molecules (diamines such as putrescine or cadaverine), the fluorescence of anthracene is “switched-off” by photo-induced electron-transfer; when a diamine is bound, however, fluorescence intensifies substantially (Fig. 3).

2 Optical sensor array concepts

The olfactory system permits differentiation among a huge numbers of chemical compounds and complex mixtures over an enormous range of concentrations. This kind of molecular recognition could not utilize the usual model of biospecificity, i.e., the lock-and-key mechanism of enzyme–substrate interaction. The olfactory receptors represent the exact opposite of that kind of specificity and show highly cross-reactive, non-specific interactions with odorants. Molecular recognition instead occurs through the pattern of response from hundreds of different types of olfactory receptor epithelia cells (each of which expresses only a single one of the hundreds of olfactory receptors found in our genome), as analyzed by the olfactory bulb and the brain.

2.1 Importance of intermolecular interactions

In applying this concept of cross-reactive pattern recognition to artificial chemical sensing, the importance of intermolecular interactions become predominant. Fundamentally, chemical sensing is molecular recognition, and molecular recognition is the consequence of interactions between molecules. The classification and strengths of inter-molecular interactions are well established (Fig. 4) and form a complete continuum from the weakest of interactions that are manifest only near 0 K to the strongest of covalent or ionic bonds. There is a seamless range from bond formation and ligand coordination, electrostatic ion–ion and proton acid–base interactions, hydrogen-bonding, halogen bonding, charge-transfer and π–π molecular complexation, dipolar and multipolar interactions, and van der Waals interactions (e.g., physical adsorption).

Remarkably, nearly all prior electronic nose technology relies essentially exclusively on van der Waals and physical adsorption, the weakest and least selective of forces between molecules. As we will argue here, colorimetric sensor arrays provide a successful method of dealing with the dilemma of sensor sensitivity vs. robustness. In many ways, colorimetric sensor arrays revisit the earlier, pre-electronic era of analytical chemistry, updated by the addition of modern digital imaging easily quantified by digital imaging and pattern recognition techniques (discussed in Section 3).

The advantage of stronger interactions for sensor arrays is both the obvious one of greater inherent sensitivity and the more subtle one of great chemical specificity. Ligation of Lewis base analytes (e.g. amines, thiols, etc.) gives bond enthalpies from ~40 to ~200 kJ mol⁻¹. In contrast, the enthalpy of physical adsorption of analytes (e.g., onto metal oxide surfaces) or absorption (e.g., into polymers) is only ~5 to 20 kJ mol⁻¹. The effective equilibrium constant for physical adsorption will typically be only about 5 × 10⁻⁵ as large as that for ligation to metal ions. Even more importantly, stronger interactions bring a much wider range of chemical interactions than simple physisorption, and consequently one may access a much higher dimensionality and improve one’s ability to discriminate among very similar analytes or complex mixtures of analytes.

Based on their recognition elements properties, the sensors used in an array will span a range of molecular specificity. At one end, there are individual sensors that are nearly completely promiscuous, i.e., highly cross-reactive; these include polymers.
and polymer blends with optical reporters embedded that adsorb analytes primarily based on hydrophobicity.\textsuperscript{117,118} Promiscuous sensors can contribute to the sum of an array’s response, but are insufficient in and of themselves to provide the differential selectivity most desirable for chemical sensor arrays. At the other extreme, there are highly selective artificial receptors that are specific for one or perhaps one closely related class of analytes. While this class of sensor can produce high specificity for specific analytes, alone they too will not make a sensor array capable of dealing with a wide range of analytes and mixtures.\textsuperscript{119} Furthermore, the synthesis of such selective artificial receptors can be complex and problematic. The optimal optical sensor array will therefore incorporate a range of colorimetric or fluorometric sensors from the promiscuous to the monogamous.

One sees exactly this range of receptors and receptor response in the olfactory system. The nose is a sensitive array of sensors able to distinguish many types of volatile analytes, but it is not equally sensitive to all analytes. The limits of detection of human olfaction are well known\textsuperscript{120} and span a range of more than 10\textsuperscript{9}. Consider the human olfactory detection limits for the simplest of analytes, those with one methyl group and one other functionality, as shown in Fig. 5. Olfactory response to methane thiol is over 1 million times stronger than that of methanol, methylamine is bound more than 100000 times more strongly than methanol, and the range from ethane to methane thiol is 10\textsuperscript{7}. Van der Waals forces, hydrogen bonding, and steric effects cannot account for such a large range for comparably sized molecules. It has been suggested\textsuperscript{121} that the olfactory receptors (ORs) are, in large fraction, metalloproteins, and that metal ion (Cu\textsuperscript{2+}, Zn\textsuperscript{2+}, etc.) ligation of strong Lewis base analytes (e.g., thiols, amines, phosphines, carboxylic acids, etc.) likely contributes substantially to the binding of many analytes by the ORs. In fact, Suslick and coworkers\textsuperscript{121} discovered a highly conserved tripod binding site in roughly 70\% of all ORs sequenced, and a very recent report from Matsunami and coworkers\textsuperscript{122} confirms the crucial role of copper ions in at least one mouse OR.

2.2 Design requirements of an optical sensor array

There are two fundamental design requirements for a colorimetric or fluorometric sensor array: (1) the chemo-responsive dye or fluorophore must contain a center to interact with analytes, and (2) this interaction center must be strongly coupled to an intense chromophore or fluorophore. The first requirement implies that it would be highly advantageous for the interaction to be more than simple physical adsorption and involve other, stronger chemical interactions.

Based on the significant intermolecular interactions responsible for optical changes, one may divide chemoresponsive dyes into roughly five separate (albeit slightly overlapping) classes: (1) Lewis acid–base dyes (i.e., metal ion containing dyes), (2) Brønsted acidic or basic dyes (i.e., pH indicators), (3) dyes with large permanent dipoles (i.e., zwitterionic, solvatochromic, or vapochromic dyes) for detection of local polarity and hydrogen bonding, (4) redox responsive dyes, and (5) chromogenic aggregative colorants (including simple ionic sulfides and plasmonic nanoparticle precursors). In addition, one may also consider environmental modifications to provide shape selectivity, either by modifying the dye peripheral superstructure or by molecularly imprinted polymers. The original colorimetric sensor array\textsuperscript{17,123,124} made use of porphyrins and metalloporphyrins as sensors, utilizing primarily aspects of Lewis and Brønsted acid–base dyes; as discussed later, the range of sensors has broadened significantly over the past decade. An example of a 36 spot sensor array for use with gas phase analytes is shown in Fig. 6.

For gas phase sensing, a colorimetric sensor array is simply digitally imaged before and during exposure to any volatile analyte, odorant, or complex mixture of odorants. The imaging is mostly commonly achieved with an ordinary flatbed scanner, but one may also use digital cameras, portable handheld readers, and even cell phones; constancy of illumination is, of course, important.
From the digital images, a difference map (Fig. 7) is easily generated by digital subtraction, pixel by pixel, of the image of the array before and after exposure: red value after exposure minus red value before, green minus green, blue minus blue. Averaging of the centers of the spots (typically ~200 pixels) avoids artifacts from non-uniformity of the dye spots, especially at their edges. The other advantage of using the differences in RGB colors is that it tends to cancel out discrepancies in printing because the color differences are only a weak function of variation of the dye concentration or spot intensity from array to array.114

The resulting data is inherently digital (simply a vector of $3N$ dimensions where $N =$ total number of spots) and all quantitative and statistically analysis is done directly from the digital difference vectors. The color difference maps are useful primarily for convenient visualization of color changes of the dye array; note that the color values are the absolute values of the differences and that expansion of the color space is useful for visualization. Note also that color difference maps shown throughout this review are generally from different arrays and therefore should only be compared within a single figure.

The choice of the individual sensor dyes in an optical sensor array is governed empirically by its intended use. One must consider if this array is meant for a broad range of analyte detection and discrimination or will it have a more specialized application. Keep in mind that the great power of optical sensors is their ability to probe the chemical properties of analytes through intermolecular interactions other than physical adsorption. If one uses only optical probes that measure local polarity (e.g., solvatochromic or vapochromic fluorescent probes doped into various polymers) then one has lost this opportunity. Potential analytes vary in their chemical properties: hydrophilicity, solubility, redox, hydrogen bonding, Lewis donor–acceptor, and proton acidity and basicity of target analytes need to be considered. In general, an optimal sensor array for general sensing purposes will incorporate as much chemical diversity among the individual sensors as possible. Given the likelihood of metal ion binding sites in the olfactory system itself,121,122 incorporation of metal ion-containing dyes into optical sensor arrays can make an important contribution to construction of a chemo-responsive sensor array. One must also consider possible interferents presented by ambient, complex environments. Finally, the stability of the dyes used in the array and their quantitative magnitude of response must of course be considered.

### 2.3 Classes of colorimetric and fluorometric sensors

#### 2.3.1 Lewis acid–base dyes (i.e., metal ion containing dyes)

**Lewis acid dyes.** Most strongly odiferous compounds are Lewis bases: thiols, phosphines, amines, carboxylic acids. Not coincidentally, these are also among the most common volatile metabolites of microorganisms; arguably, the primary function of the olfactory system is to keep us (and our digestive system) away from high concentrations of bacteria and other microbes, and hence the location of our nostrils immediately above the mouth! If one desires a sensor for the detection of such Lewis bases, then Lewis acids are the obvious solution, consistent with the likely involvment of metal ions in the olfactory system itself.121,122

Among Lewis acid dyes, metalloporphyrins (with different metals and different peripheral substituents) are a natural choice for the detection of metal-ligating vapors because of their open coordination sites for axial ligation, their large spectral shifts upon ligand binding, and their intense coloration. Indeed, the difference in color of scarlet red arterial blood and the purple of venous blood is an example of the colorimetric detection of dioxygen as it ligates to a metalloporphyrin (i.e., the iron heme of hemoglobin). In addition, it is well recognized that porphyrins show significant solvatochromic effects resulting in distinguishable colorimetric changes before and after interactions with a wide range of both ligating volatile organic compounds (VOCs) (e.g., amines, thiols, phosphines,
phosphites, thiols, etc.) and even weakly-interacting vapors (e.g., arenes, halocarbons, or ketones). Metalloporphyrins are therefore nearly ideal for colorimetric or fluorometric (for d^{10} metals primarily) detection of metal-ligating vapors. A set of representative structures is shown in Fig. 8.

**Shape and size selective metalloporphyrins.** In addition to color change, another distinguishing feature with metalloporphyrins is one’s ability to modify their periphery and provide shaped pockets to restrict access to the metal center. This capability was first developed by Collman and coworkers for selective hydroxylation of terminal methyl and methylene groups of alkanes and the selective ligation demonstrated on the first dendrimer decorated porphyrins by Moore and Suslick. This type of thermodynamic selectivity is desirable for colorimetric sensors, as equilibrium binding to shape-selective metalloporphyrins can distinguish very similar molecules from the same chemical class (e.g., branched vs. linear amines).

A relatively new class of shape-selective metalloporphyrins was developed by Sen and Suslick. A family of bis-pocketed porphyrins containing siloxyl groups on the ortho positions of a tetraphenyl porphyrin core has been shown in Fig. 9. The family contains porphyrins with six, seven, and eight tert-butyldimethylsilyl groups (denoted as Zn(Si_6PP), Zn(Si_7OHPP), and Zn(Si_8PP)), giving a set of metalloporphyrins with very similar electronic characteristics but differing steric encumbrance about the metal binding site. Zn(Si_6PP), for example, has a binding pocket of ~4 Å, greatly restricting the bonding site. These zinc complexes were sensitive to the shape and size of Lewis basic analytes; binding constants for a series of amines were found to be controllable over a range of 10^1 to 10^7 relative to Zn(TPP) and used in a colorimetric array to distinguish among alkyl amines.

**Lewis acid sensors for anion detection.** The coordination chemistry of anions was a long overlooked area of inorganic chemistry. The biological and medical importance of many anions, from the simple (Cl^−, F^−, HPO_4^{2−}, PO_4^{3−}, etc.) to the complex (ATP, lipid anions, nucleic acids, etc.), has demanded much greater attention in recent years. The supramolecular chemistry of anions also plays essential roles in catalysis and environmental sciences.

There has been tremendous recent effort in the design of anion receptors for sensing by colorimetric or fluorometric means. The use of Lewis acid dyes for the detection of anions has been an active area of research and has been extensively reviewed recently. There are, however, unique challenges to these studies because anion complexation is quite different from that of metal cations, largely because of the relatively large size of anions and the omnipresence of protons in aqueous media. Anion receptors can be neutral or positively charged and in general anion-receptor interactions are dominated by electrostatics and hydrogen bonding. It is common to link a chromogenic or fluorescent reporter moiety to a specific chelating receptor, but one may also use fluorescent Lewis acids directly.

Displacement assays, dyes with urea, thiourea, or naphthalimide sites, or metal ion containing dyes (especially of lanthanide and labile d^8 and d^{10} transition metal ions) have all been explored especially heavily as anion binding sites for both

---

**Fig. 9** Bis-pocketed zinc siloxylporphyrins for shape-selective discrimination of Lewis base analytes. Upper: chemical structures. Lower: molecular models (framework side view and space-filled top-view) of Zn(Si_6PP).

**Fig. 10** Upper: a sampling of traditional complexometric indicators. Lower: Eriochrome Black T in aqueous solution.
colorimetric and luminescent detection. Work on colorimetric and fluorometric sensors for specific anions is important, but mostly beyond the scope of this review.

**Lewis base sensors for cation detection.** Chelating and macrocyclic ligands are, by definition, Lewis bases. Modern supramolecular chemistry finds its origins in the design of crown ethers, cryptands, *etc.* and their size specific binding of metal ions.\(^{146}\) However, the use of semi-specific chelating Lewis acid dyes for colorimetric sensors of metal ions, so-called complexometric indicators\(^{147,148}\) (Fig. 10), dates back more than 150 years.

Complexometric indicators are used to chelate metal ions while simultaneously inducing a color change. These chromogenic or ionochromic dyes are designed to bring about a specific color change in the interaction with metal cations. Classical complexometric indicators (many of which are natural products also used as histological stains and some of which date back to the early 1800s) may have greater or lesser degrees of specificity: for example, calcein and Eriochrome Black T are used to detect Ca\(^{2+}\), Mg\(^{2+}\), and Al\(^{3+}\); hematoxylin for Fe\(^{3+}\) and Al\(^{3+}\); murexide for Ca\(^{2+}\), Cu\(^{2+}\), Ni\(^{2+}\), and rare earth ions; and xylene orange for Ga\(^{3+}\), In\(^{3+}\), and Sc\(^{3+}\). Traditionally, complexometric titrations were displacement reactions, starting with the metal ions bound to the indicator and then displaced by the addition of EDTA, so that the free dye (rather than the metal ion complex) served as the endpoint indicator. Recent interest in metal ion sensors have taken advantage of the cross-reactivity of complexometric indicators to generate solution based arrays, for example using a microtiter plate or an immersed membrane\(^{149–152}\) for simple identification of single metal ions in water.

2.3.2 **Bronsted acidic or basic dyes (i.e., pH indicators).**

The origins of chemistry as a discipline are closely tied to our fascination with “pretty colors” and the importance of the dye industry to early chemists can hardly be overstated.\(^{153–155}\) Many dyes, of course, change their colors depending on the pH. Litmus (7-hydroxyphenoxazone) was available even to alchemists in the Medieval times and literally means “colored moss” in Old Norse (litmus is produced by lichens, particularly *Roccella tinctoria*). There are, of course, dozens of pH indicators derived from natural products, especially the anthocyanin oxonium dyes from blueberries to grapes to red cabbage to rhubarb (Fig. 11).

Synthetic pH indicators received enormous effort during the first half of the 20th century,\(^{113}\) but even today substantial interest in new formulations (including sol–gel encapsulated indicators and indicators suitable for intra-cellular use in *in vivo*) continues.\(^{99,156,157}\) An immense variety of organic chromophores (e.g. azo dyes, nitrophenols, phthaleins, sulfophthaleins, anilinesulfophthaleins, triphenylmethane dyes, *etc.*) were created largely to measure the pH of aqueous solutions or as histological stains for biomedical applications.\(^{113,158}\) The \(pK_a\) values among various pH indicators for aqueous solutions range, of course, from below 0 to 14. The chemical diversity of some pH indicator dyes that have been used in colorimetric sensor arrays is shown in Fig. 12.

2.3.3 **Solvatochromic and vapochromic dyes.** Dyes whose dipole moments are significantly different between their ground and excited states will show color changes depending upon the polarity of their environment: *i.e.*, solvatochromism. If the excited state has a larger dipole moment, it will be more stabilized relative to the ground state in a more polar environment and *vice versa.* Nearly all dyes inherently show some solvatochromism. For “solvatochromic” dyes, these changes in dipole moments are very large, leading to impressive color or

---

**Fig. 11** Some naturally occurring pH indicators. Litmus is from a lichen, delphinidin from cabernet sauvignon, and cyanidin from blueberries.

**Fig. 12** A sampling of representative pH indicator dyes.

**Fig. 13** Brooker’s merocyanine dye, (1-methyl-4-[(oxocyclohexadienylidene)-ethylidene]-1,4-dihydropyridine), and its solvatochromic shifts in various solvents.
fluorescence changes that depend upon the polarity of the solvent in which the dye is dissolved (Fig. 13).

Common classes of solvatochromic dyes include the merocyanines, azobenzenes, oxazines, thiazines, nitro-amino-substituted polythiophenes, and pyridinium N-phenolate betaine dyes. A common feature of most solvatochromic dyes is that they are “push–pull” systems (Fig. 11) with a strong zwitterionic component to their electronic structure, i.e., a large conjugated π system with strong electron donor groups at one end and strong electron withdrawing groups at the other.

Several solvent polarity scales159 have been based on the wavelength shifts of optical transitions of solvatochromic dyes, including Kosower’s and Brooker’s early studies, Reichardt’s ET159–163 and Taft’s π*164,165 scales. Solvent polarity is very much a multi-parameter property, involving dipolar, quadrupolar, and multi-polar interactions, hydrogen bonding donation and acceptor properties, Lewis acid–base interactions, etc. Thermodynamic and theoretical analysis of the origin and meaning of solvation, solvent polarity, and solvatochromism continue actively,166–169 especially useful are thorough comparisons among all related multiparameter descriptions of solvent polarity.159,167,170,171 Of particular interest for optical sensor arrays, the use of polymers doped with a solvatochromic dye (often Nile red) as a reporter on swelling of the polymer by absorbed analytes has been heavily used for optical fiber sensors.6,172,173

A separate class of solid-state materials that provide a colorimetric response to solvent vapors are referred to as vapochromic or vapoluminescent solids.174,175 These are most commonly porous coordination complexes, particularly of square planar Pt(II) d8 compounds. The vapochromism is triggered by intercalation of solvent molecules into the porous crystals and the color and luminescence changes derive from changes within the solids from weak interactions including coordination of solvent molecules to the metal centers, metalophilic contacts, π–π stacking, hydrogen bonding, and general non-specific host–guest interactions; these interactions can lead to changes in the ordering of excited states, leading to large luminescent differences. A particularly striking example has been recently reported using triarylboron-functionalized phenylacetylide platinum(n) square planar complexes.176 Chiral vapochromic materials have also been used for identification of enantiomeric vapors.177 Because vapochromism requires intercalation into the interstices of crystalline materials, the response time can be somewhat slow (although controllable, perhaps, using nanocrystalline morphologies), and the weak interactions responsible can also lead to limited sensitivity of these materials as sensors.

2.3.4 Redox indicator dyes. Oxidation/reduction (redox) indicators are colorimetric reagents which show a distinct color change at a specific electrode potentials. These are all organic compounds exhibiting reversible redox reactions. Examples include anilinic acid, diphenylamine, eriogreen, m-cresol-indophenol, methylene blue, and nile blue.178–180

Due to the fact that the majority of redox indicators engage a proton as a participant in their electrochemical reactions, redox indicators are sometimes divided into two groups separating those that are dependent on pH and those that are not. In order to make an optical sensor for sensitive detection of hydrogen peroxide (i.e. in the concentration range of 10−8 to 10−1 mol L−1), a redox indicator (e.g. Meldola blue) was used in a sol–gel layer.180

In order to have sensitive detection of triacetone triperoxide (TATP) vapor (i.e. in the range of 50 ppb to 1 ppm), a colorimetric sensor array based on redox dyes including Lissamine Green B, o-dianisidine, diphenyl amine, N-phenyl-1,4-phenylenediamine and N,N'-diphenyl-1,4-diphenyldiamine (Fig. 14) was constructed by Lin and Suslick.181 Using a method of hydrolyzing TATP vapor to constituent acetone and hydrogen peroxide, the array was capable of detecting concentrations of TATP vapor down to 2 ppb.

2.3.5 Chromogenic aggregative colorants. Chromogenic colorants whose color is altered by aggregative phenomena has become an area of massive recent development, especially for biosensing.182–184 Processes that cause aggregation, dispersion, or formation of colloidal materials generate changes in color and fluorescence through various mechanisms, ranging from simple absorbance and scattering by colloidal solids to plasmonic absorbance to quenching of attached or adsorbed fluorophores. The simple precipitation of metal salts or formation of metal nanoclusters on reaction with thiols and sulfides goes back to the earliest qualitative spot tests.111,112

The recent cutting edge is represented by control over the nanostructure of optical sensors. For example, gold nanoparticle (NP) agglomerates that efficiently quench adsorbed fluorophores; analyte binding can disperse such nanoparticle agglomerates and create a fluorescence turn-on. Judicious choice of NP functionalization and of fluorophore provides a versatile platform for solution phase sensing.

2.3.6 Displacement strategies for fluorescent probes. There are three general classes of fluorescent indicators for chemical sensing in solution: (1) intrinsic probes (where the sensor is itself fluorescent), (2) conjugated or extrinsic probes (where a fluorophore is conjugated to the sensor binding site and its fluorescent properties modulated by analyte binding), and (3); displacement, dissociation, or differential probes (where the analyte competitively binds to an artificial or natural receptor that also binds a fluorophore).99,102,103,185
Displacement probe strategies require some sort of reversible interaction between a receptor (natural or artificial) and a reporting fluorophore/chromophore; the binding of the reporter molecule must modify the fluorescence or color the reporter. The receptor can be either specific for one class of analytes or more cross-reactive.

Indicator displacement assays (IDA) use a parallel set of multiple not-too-selective displacement probes, often referred to as differential selectivity, to generate a pattern of response not unlike the colorimetric sensor arrays already discussed. This approach has been particularly well explored by the Ansglyn group.181–183 The potential disadvantage of displacement strategies, of course, is diminished sensitivity because there is an inherent competition between the analyte and the already bound fluorescent or colorimetric reporter.

One normally thinks of sensor arrays as single physical solid devices: a printed array on a polymer membrane or a bundle of fiber optic probes, etc. Displacement strategies, however, are generally limited to solution phase sensing and do not lend themselves easily to a solid-state sensor array platform that could be immersed in a solution of analytes. Instead, solution phase array sensing is carried out by a parallel analysis of multiple aliquots of the analyte solution, each with an added, different homogeneous probe, e.g. using microwell plates with a microwell fluorescent scanner.

Kubarych, Adams, and Ansglyn used a set of commonly available proteins with a set of fluorophores as non-specific probes for hydrophobic molecules.186 In this case, non-specific hydrophobic binding interactions were used with an indicator displacement strategy to provide multiple diverse probe sites within a single protein molecule, as shown schematically in Fig. 15. This led to the ability to differentiate among several different hydrophobic species including fatty acids and food oils.

2.3.7 Molecularly imprinted optical sensors. One way to improve optical sensor resistance to interferents is the application of a molecular imprinting technique.187 The molecular imprinting is a process which can rapidly synthesize polymers with differential selectivity to targeted analytes, which include both molecular and ionic species as well as enantiomers.188–195 In general terms, one includes a non-polymerizing analyte in the monomer solution during polymerization. The templating analyte is then exhaustively removed from the resulting polymer. This molecularly imprinted polymer (MIP) must be sufficiently crosslinked to retain internal structural integrity, but not so rigid as to prevent template removal. As with displacement assays, MIP sensors face the problems of specificity and are also generally limited to solution phase sensing.

As an example, Shimizu and coworkers189 constructed a sensitive, selective receptor array with a molecular imprinting procedure which was able to classify different amines (Fig. 16). They applied a dye-displacement strategy for rapid and versatile measurable colorimetric response. The presence of molecularly imprinted polymers (MIP) with high cross-reactivity as detection elements provided a specific response for each analyte.

2.4 Substrate considerations

2.4.1 Printed arrays. While the choice of chemoresponsive dye or fluorophore will dominate the effectiveness of any optical sensor array, the functionality of the array will also be influenced by the substrate and morphology of the substrate upon which the colorant is placed. Sensitivity, reliability, accuracy, response time, susceptibility to interferents, and shelf-life of the array can be heavily influenced.

A wide variety of solid supports have been used for colorimetric and fluorometric array construction. The desired properties of such substrates include inertness towards gases and liquids, high surface area (to incorporate sufficient colorant), optical transparency or high reflectivity, and stability over a wide pH range.14,16,48,85,89,156 A simple method for array manufacture involves printing dye formulations on the surface of reverse phase silica gel plates, acid-free paper, or porous polymer membranes made out of a material such as cellulose acetate or polyvinylidene difluoride (PVDF). Robotic pin printers serve this function particularly well (Fig. 17). Ink-jet printing works well for a limited number of sensors, and spin-coating can also be used, but is difficult for multiple spot arrays.

A major shortcoming of many chemical sensors is their sensitivity to changes in ambient humidity.6,8 For any real world application, changes in humidity from day to day or from indoors to outdoors can involve changes of tens of thousands of ppm in water vapor concentration. Any significant response to humidity...
can prove lethal for detection of ppm concentrations of VOCs. The use of hydrophobic materials (e.g., PVDF membranes) as colorimetric or fluorometric sensor array substrates can be highly advantageous in reducing array response to humidity over a wide range (10 to >95% relative humidity)\(^\text{14,17,114,197,198}\) and can even permit the use of such arrays in aqueous solutions.\(^\text{199}\)

For sensing applications, soluble, molecularly-based dyes are deposited as a viscous film or on a high surface area membrane or solid, and the analytes can gain access to the colorant with an acceptable response time. In contrast, pigments (which by definition are insoluble colorants) are not generally permeable to analytes and therefore reactive only on their outermost surfaces, which dramatically reduces any colorimetric or fluorometric response to the presence of analytes. Molecular dyes, however, often have limited shelf-life as sensors, particularly when incorporated into viscous films or polymers due to crystallization (and consequent loss of analyte accessibility to the colorant centers).\(^\text{200}\)

Porous sol–gel glasses can also provide excellent matrices for chemically responsive colorants.\(^\text{85,156,201,202}\) An effective nanoporous pigment sensor can be made by adding chemoresponsive dyes to ormosils prepared from suitable silane precursors in low volatility solvents.\(^\text{16,197,200,203,204}\) In addition, the physical and chemical properties of the matrix (e.g., hydrophobicity, porosity) can be easily altered using organically modified sol–gel (i.e., “ormosil”) formulations, and different silane precursors are needed depending on the solubility of the dye. The use of these porous pigments significantly improves the stability and shelf-life of the colorimetric sensor arrays and permits direct printing onto non-permeable polymer surfaces.\(^\text{16,200,201,205}\) Additionally, we have observed that the matrix may serve as a preconcentrator, improving the overall sensitivity.

Digital image analysis shows that printing of these nanoporous pigments have a uniform color distribution across the center of the spot (Fig. 18).\(^\text{200}\) Reproducibility of the optical densities of printed spots is excellent, and chemical sensing experiments generally use the difference between before-exposure and during-exposure image, which further reduces errors in the pattern analysis. It is also important that image analysis of printed spots utilize the average RGB values of the center portion of the printed spots to eliminate artifacts from the spot edges.\(^\text{114}\)

The development of optically based chemical sensing platforms has increasingly employed substrates manufactured with

---

**Fig. 17** A robotic pin printer (upper) with a slotted pin printing array (lower) is capable of printing hundreds of sensor arrays per hour.

**Fig. 18** (a) Image of a printed array, 2.5 × 2.5 cm. (b) RGB values of a linescan across the centre of a typical spot. Used with permission from ref. 200.

**Fig. 19** SEM micrographs of a 1 mm diameter spot of a porous ormosil pigment printed on PET film: (a) top surface and (b) energy dispersive spectroscopic (EDS) elemental mapping (Si Kα); (c) cross-section and (d) EDS elemental mapping (Si Kα). Used with permission from ref. 200.
Advanced processing techniques, and the control of over sensor morphology, micro- and nano-structure must be characterized.\textsuperscript{21} Scanning electron microscopy has shown that the nanoporous pigment films are typically \( \sim 3-4 \) \( \mu \text{m} \) thick with uniform silicon content throughout the spot, and transmission electron micrographs reveal the pore structure of these ormosil films (Fig. 19 and 20), which assist in mass-transport, and are responsible for the fast response times observed during sensing experiments (90% response generally occurs in \(< 2 \) minutes).\textsuperscript{200}

Another means of generating nanoporous pigments has been reported. Bang et al.\textsuperscript{205} prepared silica microspheres using tetra-methoxysilane (TMOS) and methyltrimethoxysilane precursors (MTMS) in which chemoresponsive pH and solvatochromic dyes were incorporated with an ultrasonic-spray aerosol–gel synthesis method.\textsuperscript{206}

### 2.4.2 Fiber optic arrays

A powerful alternative, especially for fluorescent sensors, is the use of fiber optic arrays. Almost twenty years ago, Walt and co-workers began developing multi-fiber optical bundles as cross-reactive or multi-receptor high-density sensor arrays.\textsuperscript{207} In this technique, bundles of very small silica fiber-optic cables are chemically etched to create a 2-dimensional array of microwells, which are used to hold a random distribution of beads containing individual fluorescent probes. In order to ameliorate the loss of physical position as a useful measurable property, these beads are functionalized with optical encoding elements so as to identify themselves during the analysis. A graphical explanation of this process is shown in Fig. 21 and SEM images of the final arrays are shown in Fig. 22.

For achieving simultaneous, multi-analyte, high-density, and high-throughput sensing analysis, the optical fiber-based arrays were developed using optical fiber bundles that comprise thousands of individual single-core fibers which are individually modified with a diverse sensing chemistry using a random assembly method. In general, the fiber arrays (with total size of around 1 mm) contain a few thousand up to a hundred thousand individual fibers with size of 2–10 \( \mu \text{m} \) prepared by selective etching of the polished array in acid solution\textsuperscript{208} (Fig. 22). This engineered structure enables individual wells to be recognized by the optical fiber defining its base, providing a high-density array of micro-wells that can be simultaneously and independently interrogated by light. Excitation light is introduced into the unfunctionalized end of the fiber; emission signals from individual sensors return through the fiber and are magnified and projected onto a charge-coupled device (CCD) camera, leading to the simultaneous observation of all sensors.\textsuperscript{209}

These fiber-based fluorescent sensors have found impressive use in biosensing for immobilization of desired bio-molecular sensors,\textsuperscript{172,211,212} single molecule detection,\textsuperscript{213,214} and even whole cells\textsuperscript{215,216} on the fibers.\textsuperscript{217,218} Living cells and beads containing biological recognition elements are loaded into the micro-wells (with the volume of femtoliter).\textsuperscript{209,219,220}

### 3 Statistical analysis and modeling

Chemical property space has a very high dimensionality because of the large number of different chemical properties that are

---

**Fig. 20** TEM micrograph of porous ormosil pigment showing the 50 to 200 nm pore structure created in these ormosil xerogels. Used with permission from ref. 200.

**Fig 21** General strategy used for creating bead microarrays using etched fiber-optic bundles. Types A, B, and C (blue, green, and red respectively) indicate the types of self-encoded beads, while the insets represent a typical time-based sensor response to an exposed analyte. Reproduced with permission from ref. 210.

**Fig. 22** SEM images of etched fiber-optic bundles. (a) Optical fiber-based array with hexagonal pits contains 50,000–60,000 fibers; inset shows magnified views of the individual fibers (bright dots); (b) microwells etched into the individual fibers and surrounded by cladding; darker gray circles correspond to the ends of the optical fibers defining the bottom of each well; (c) beads loaded into wells. Reproduced with permission from ref. 173.
largely orthogonal to one another (e.g., Lewis acid-base, Bronsted acid-base, redox, electrophilicity, nucleophilicity, hydrogen bonding, polarity, etc.). To differentiate among all possible volatile compounds and the huge assembly of possible mixtures of those compounds requires highly multidimensional data representing a wide range of chemical properties space. Fundamentally, this is why the olfactory system evolved to incorporate hundreds of highly cross-reactive receptors: the combinatorials of the responses of those cross-reactive sensors provides the high dimensional data used by the olfactory bulb for pattern recognition of odors.

If one bases a sensor array on physical properties, however, there is a very limited dimensionality to the resulting data. Physical adsorption–absorption will dominate any interaction of analytes either with simple surfaces (e.g., metal oxide electrical sensors or chemFETs, etc.) or with polymer coatings (e.g., coated quartz microbalances, conductive or composite polymer sensors, etc.). The primary contributions to physisorption (van der Waals, weak hydrogen bonding, polar interactions) are roughly equivalent to what chemists refer to as “hydrophobicity”.

As a consequence, one of the dirty little secrets of electronic noses is that they are only rarely multidimensional arrays in a statistical sense.\textsuperscript{5,8,18} The sensor may be a physical array (e.g., an array of ten or twenty different metal oxides or a dozen different conductive polymer composites), but that does not in and of itself make the sensor a multidimensional array for analytical or statistical purposes. In general, in such physical arrays, there is in fact only one overwhelming dominant dimension that contains >90% of the total variance among analytes: most often, that dimension is essentially hydrophobicity. As a consequence of this intrinsic low dimensionality, most conventional electronic nose technology is not able to distinguish among large libraries of similar complex mixtures.

Low dimensional data does have some advantages. Statistically, analysis is simplified, but often at the cost of more limited discriminatory abilities. The primary advantage of relying on weak analyte–sensor interactions such as physical adsorption is improved reversibility, especially over short periods of time, but even this is a two edged-sword: the reversibility comes directly from the weakness of the interaction, which implies diminished sensor sensitivity!

In contrast, sensor arrays based on chemical properties have intrinsically a much higher dimensionality. Having a high dimensionality has the advantage of much greater ability, at least in principle, of being able to differentiate among analytes with much greater discriminatory power. The disadvantage of limited reversibility for strong interactions can be overcome by making the sensor array disposable, as we have already discussed, and is counterbalanced by the improvement in sensitivity and improved limits of detection.

The greater dimensionality, however, must also involve a more sophisticated approach to statistics than that which chemists are often comfortable.\textsuperscript{221} The inherent problem with high dimensionality is that the analytical volume increases much more rapidly than the available data, so the datasets are often formally “sparse” compared to the total size of the parameter space. In addition, this “curse of dimensionality” creates difficulties for function approximation, model fitting, information extraction, as well as computation.\textsuperscript{222} Statistic methods for multidimensional data all share the common goals of displaying multidimensional data effectively, evaluating data sets, and predicting the identity of unidentified samples based on a known library.

There are a variety of statistical methods available to deal with high dimensional data well beyond the scope of this review.\textsuperscript{221,223} We will give an overview here only of the three most common approaches: hierarchical cluster analysis (HCA), principal component analysis (PCA), and linear discriminant analysis (LDA).

In general, for chemometric data there are two distinct statistical approaches: clustering vs. classification.\textsuperscript{223–225} Cluster analysis essentially tells one what resembles what, e.g., how close the vectors representing data are to one another in a high dimensional space. Classification analysis, on the other hand, attempts to predict to which category (among a fixed number of known categories) any particular (new) datum belongs.

Statistical methods can be either biased, in which case the evaluation algorithm is told of the class identities of individual cases, or unbiased (or model-free), where all cases are evaluated identically regardless of class identity. Unbiased methods are typically used to evaluate a data set to provide a semi-quantitative idea of the quality of the data set and follow simple, straightforward algorithms. Biased methods, on the other hand, can provide significantly more power and utility with a concomitant increase in complexity, but at the cost of demanding datasets for which one already knows the answers. Biased methods can be predictive, allowing for class assignment of new experimental cases by using a training set.

### 3.1 Hierarchical cluster analysis (HCA)

HCA is an agglomerative clustering technique whereby clusters are determined from the Euclidean distance between experimental data. In its simplest form, nearest-neighbor points are paired into a single cluster which is then paired with other nearest-neighbor points or clusters until all points and clusters are connected to each other, shown schematically in Fig. 23.\textsuperscript{224,225} The most common clustering criterion used in HCA is Ward’s

![Fig. 23](image-url)
minimum variance method, which minimizes the total within-cluster variance.

The resultant dendrogram shows connectivity and some measure of the distance between each of the pairs. In the context of chemical analyses, these two important pieces of data answer two questions: connectivity explains relationship similarity, i.e. ‘what species/samples are similar to each other?’ and distance explains magnitude, i.e. ‘how similar are they?’.

There are three primary limitations to the HCA technique. The first involves fundamental limitations of all unbiased methods: HCA is not easily capable of predictive analysis. Second, dendrograms created using HCA must be re-created with each addition of a new analyte, so comparing dendrograms (even with a very similar data set) is typically only useful for rough qualitative purposes, i.e. ‘what does this new sample look most like?’. The third limitation is that of interpretation of noisy data. One must be cognizant that the dendrograms are essentially “mobiles” and that rotations around clustering axes do not represent meaningful differences between dendrograms, as shown in Fig. 24: mis-clustering of noisy data can be easily misinterpreted.

Despite these limitations, dendrograms provide a straightforward method of displaying cluster similarity semi-quantitatively. A representative example containing 100 VOCs is shown as Fig. 25; it is worth noting that similar chemical classes cluster tightly together, which is a consequence of the reactivity of the particular sensor array.

3.2 Principal component analysis (PCA)

PCA is a dimensional reduction technique that condenses the variance among several possibly-correlated dimensions by creating a new orthogonal set of dimensions using linear combinations of the initial dimensions. These new dimensions (also called directions, components, etc.) are ranked such that the first dimension explains the largest amount of data variance, the second dimension explains the second largest, and so on. One typically seeks a number of new orthogonal dimensions sufficient to encompass at least 95% of the variance. Plots using the resulting set of principal components are often easier to visualize than the original data set, but only if the original dataset is actually low dimensional in a statistical sense. As discussed earlier, electronic nose data often requires only two or perhaps three principal components to express the true variability among the data, regardless of the number of different sensors in the physical array. For low dimensional data, PCA therefore provides a straightforward method of displaying sample set variability, i.e. ‘how similar are these species/samples to each other?’.

As we shall see in Section 3, when dealing with a large number of analyte classes, a sensor array designed to probe a large reactivity space (i.e., an array with high dimensionality in a statistical sense) is highly desirable. If one is examining a narrow class of analytes, however, then apparent high dimensionality of a sensor array, regardless of the number of actual physical sensors, becomes indicative of large amounts of noise relative to total variance: in such cases, in the absence of noise, the theoretical maximum dimensionality is equal to the lesser
of either the number of sample classes or the size of the chemical reactivity space.

As a consequence, the dimensionality of one’s data is not determined directly by the number of different sensors in one’s array. As an example, Lin and Suslick\textsuperscript{181} designed a colorimetric array containing 16 redox-sensitive dye formulations to detect specifically strong oxidants and peroxy-based explosives and was not designed with other reactivity properties in mind. This array probes only a small chemical reactivity space, and the PCA reveals that only two dimensions (n.b., not 16) were required to reach 95% variance, as shown in Fig. 26.

Since principal components are combinations of individual array component responses, the number of dimensions required for 95% variance provides information about the range of analyte-sensor interactions being probed; an array that probes only pH, for example, may only have one dimension required to reach 95% variance, while an array that probes pH, hydrophobicity, dipole moments, film permeability, and nucleophilicity can be expected to require at least five dimensions. PCA is thus a powerful tool for evaluating sensor arrays, especially those with multiple disparate components, as it allows some insight into the sensor’s “chemical reactivity space”, i.e. the number and possibly identities of chemical interactions being probed by the sensor array. A scree plot, showing the cumulative contributions of each principal component, provides a quantitative measure of the contributions of different orthogonal reactivities to the variance of the array response (Fig. 27).

Like HCA, PCA is an unbiased method that is best suited for evaluation of data sets rather than prediction. PCA, however, can make rudimentary prediction methods possible, especially if the data set is low dimensional and has a large separation among sample classes. If the data set does not have a large separation, however, PCA may not adequately be able to predict the identity of an experimental sample. Examples of these are shown as Fig. 28.

3.3 Linear discriminant analysis (LDA)

Like PCA, linear discriminant analysis (LDA) is a dimensional reduction technique that constructs a set of orthogonal dimensions used to describe the data; LDA, however, seeks to find a set of dimensions that best separates data into already known classes, rather than simply describing the total variance. Unlike HCA or PCA, LDA is a biased method; statistical analysis using LDA requires inputting a class label for each sample. Components of each dimension are ranked in order to maximize the ratio of between-sample variance to within-sample variance; i.e. it ranks components based on their signal to noise ratio as compared among differing sample classes.

LDA can be used to predict the identity of unknown samples by using a training set, similar to PCA. However, because the dimensional components are optimized to maximize differentiability, LDA will show better ability to differentiate among sample classes. A general example of this improvement is shown as Fig. 29.

The primary weakness of LDA is related to sample size. All statistical methods require multiple observations in order to determine any useful data (e.g. mean, variance, etc.); LDA is
unique among the three methods presented here, however, as sample class covariances must also be determined in order to allow for comparison among classes. Because of this, the covariance matrix tends to be unstable when sample size is not significantly larger than the number of sample classes being analyzed, and this is more problematic for high-dimensional data. Consequently, LDA can give drastically fluctuating results with small sample sizes (compare to PCA or LDA, which can be unreliable with small sample size, but not unstable). A representative example of a two-component LDA plot is shown as Fig. 30.

As an improvement on LDA, tensor discriminant analysis (TDA) is an array generalization of LDA better able to take advantage of high dimensionality. More precisely, tensor discriminant analysis is used to classify multi-way array measurements (i.e., "tensor measurements"), rather than one-way vector measurements. For example, the data collected using colorimetric sensor arrays can be viewed as a 3-way tensor with the first way corresponding to choice of the dye, the second way corresponding to the effects of the color changes (i.e., $\Delta R$, $\Delta G$, $\Delta B$, which are not fully independent for any one dye), and the third way corresponding to time progression (for kinetic responses). The general strategy of tensor discriminant analysis is to find orthogonal linear classifiers, which are essentially linear combinations of the three-way interactions of the effects of the dye spot choice, the three color changes of each spot (i.e., $\Delta R$, $\Delta G$, $\Delta B$), and the temporal evolution, to maximize the ratio of between-class variation to within-class variation. Tensor discriminant analysis can greatly improve the sensitivity, specificity, and computational efficiency of discriminant analysis method because of the dimensionality reduction. For example, if there are 36 dyes, then LDA would have to deal with $36 \times 3 = 108$ dimensions, whereas TDA would reduce that to $36 + 3 = 39$ dimensions, thus eliminating 69 dimensions (i.e., $108 - 39$).

4 Applications of optical sensor arrays

4.1 Discrimination of volatile organic compounds

The colorimetric sensor array first developed by Rakow and Suslick used an array of different metalloporphyrins exclusively for the visual identification of different families of organic vapors. Ligation of analytes to metalloporphyrins induced large color changes that were used for their identification. The sensor array was able to respond to a wide range of organic compounds such as alcohols, amines, ethers, phosphines, phosphites, thioethers, thiols, arenes, halocarbons and ketones, often with sensitivities below 1 ppm and importantly, without response to change in humidity. Using different metalloporphyrins with a wide range of chemical hardness and ligand-binding affinities as well as solvatochromic effects allowed differentiation among a wide range of volatile analytes.

By broadening the types of sensors in the colorimetric array to include shape selective bis-pocketed porphyrins, pH indicators, and solvatochromic dyes to a total of 24 sensors, Rakow et al. were able to demonstrate highly selective discrimination among very closely related amines, with sub-ppm sensitivities.
Discrimination among linear alkyl amines, and even isomeric amines was possible, as shown in Fig. 31. Similar arrays were tested by Tang et al. 231 and by Luo et al. 232 who were able to demonstrate limits of detection for trimethyl amine and ammonia, respectively, below 50 ppb.

With further expansion of the array to 36 colorimetric sensors, Suslick and coworkers 114 were able to demonstrate error-free discrimination among 100 different VOCs with common organic functionalities including primary, secondary, tertiary, and aromatic substituents of amines, arenes, alcohols, aldehydes, carboxylic acids, esters, hydrocarbons, ketones, phosphines, and thiols. The array discriminates among VOCs by probing a wide range of intermolecular interactions, including Lewis acid–base, Brønsted acid–base, metal ion coordination, hydrogen bonding, and dipolar interactions. LODs are analyte dependent and were not determined in this study, but were generally in the low ppbv range for amines, carboxylic acids, thiols, and phosphines. The sensitivity of the array to bases and acids is a result of the strong metal-analyte interactions, either by metal ligation (i.e., coordination or dative bonding) or by Bronsted acid–base interactions. Weakly coordinating vapors such as esters, ketones, alcohols, arenes, and hydrocarbons show a lower response, just as the mammalian olfactory system.

Importantly, by proper choice of dyes and substrate, the array is essentially non-responsive to changes in humidity. A selection of the difference maps of a representative subset of 24 VOCs are presented in Fig. 32.

The PCA of the dataset representing the full 100 VOCs shows an extraordinarily high level of dispersion by the colorimetric sensor array: 14 dimensions are required to define 90% of the total variance, 22 dimensions for 95% of the total variance, and 40 dimensions for 99% (Fig. 27).

Most remarkably, because the colorimetric sensor array is based on analyte-array chemical reactivity, chemical class information becomes readily available from the data analysis. 114 Fig. 32 shows the familial similarities of the color difference maps of alkyl amines vs. aromatic amines vs. carboxylic acids vs. aldehydes vs. phosphines, simply by inspection. The more detailed analysis afforded by HCA shows how well that chemical class information can be revealed (Fig. 25).

While these colorimetric sensor arrays work exceedingly well for reactive volatiles, they have not had especially high sensitivity to less reactive vapors. For example, common VOC indoor air pollutants (e.g., aromatic hydrocarbons, chlorocarbons, other organic solvents) are generally not especially reactive and are not detected at low concentrations. Lin, Jang, and Suslick recently reported 19 a dramatic improvement in the sensitivity of colorimetric sensors for the detection and identification of such less-reactive VOCs by the use of a disposable pre-oxidation technique in which the analyte-stream was passed through an oxidation tube (of chromic acid on silica) before reaching the sensor array.

Fig. 33  Left: schematic illustration of the preoxidation technique. A Teflon tube is packed with chromic acid to pretreat the gas flow containing a VOC before it is passed over the colorimetric sensor array. Center and right: responses to p-xylene at IDLH concentrations without (center) and with (right) pre-oxidation tube. Reproduced with permission from ref. 19.

Fig. 34  HCA dendrogram for 20 commonly found indoor pollutant VOCs at their IDLH concentrations and a control. All experiments were run in quintuplicate with 30 mg chromic acid on silica as the pre-oxidation reagent; no confusions or errors in classification were observed in 105 trials. Reproduced with permission from ref. 19.
array (Fig. 33). Preoxidation of VOCs produces reactive species such as carboxylic acids, phenols, and aldehydes which have a stronger interaction with the colorimetric sensor array. This results in a ~300-fold enhancement of sensitivity with a concomitant increase in discrimination ability.

Since each VOC produces a different mixture of oxidized derivatives, the array response to these more reactive volatile by-products provides a unique, but much more sensitive, signature for the initial VOC. 20 commonly found VOC pollutants in indoor air were examined as representative analytes and all were discriminable by HCA, as shown in Fig. 34, both at their immediately dangerous to life or health (IDLH) and at their permissible exposure limit (PEL).

4.2 Toxic industrial chemicals

Toxic industrial chemicals (TICs), by their very nature, are chemically reactive. The toxicities inherent in toxic industrial chemicals derive from a very wide range of specific chemical reactivities that affect multiple systems within living organisms. Some acute toxins target specific, critical metabolic enzymes (e.g., HCN inhibits cytochrome c oxidase while phosphine inhibits pulmonary function); some cause cell lysis in the lungs creating pulmonary edema (e.g., HCl, HF) and others are potent oxidants or reductants that can target various biosystems. There is an obvious need for rapid, sensitive identification, and determination of TICs, yet we have no small and inexpensive technology for personal dosimetry of TICs in the chemical workplace or by first responders to industrial fires or chemical spills.

Prior electronic nose technology, which utilizes weak analyte–sensor interactions, have had a limited ability to detect compounds at low concentrations relative to analyte saturation vapor pressure and therefore are often unable to detect TICs at their IDLH (immediately dangerous to life or health), PEL (permissible exposure limit) concentrations. In addition, interference from large environmental changes in humidity or temperature remains highly problematic.

Colorimetric sensor arrays, however, are exceptionally well designed for the detection, identification and quantification of TICs due to their reliance on analyte chemical reactivity. Suslick and coworkers in a series of papers developed the use of nanoporous sol–gel pigments for the chemoresponsive elements of an extremely sensitive colorimetric sensor array. They selected high hazard TICs from the reports of the NATO International Task Force 25 and 40 and examined the ability of their array to discriminate among the 20 TICs shown in Table 1.

The sensor array was able to discriminate without error among these 20 TICs at both their IDLH concentration within two minutes of exposure and at PEL concentration within five minutes.

### Table 1

<table>
<thead>
<tr>
<th>TIC</th>
<th>IDLH (ppm)</th>
<th>PEL (ppm)</th>
<th>Extrapolated LOD (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ammonia</td>
<td>300</td>
<td>50</td>
<td>0.08</td>
</tr>
<tr>
<td>Arsine</td>
<td>3</td>
<td>0.05</td>
<td>0.01</td>
</tr>
<tr>
<td>Chlorine</td>
<td>10</td>
<td>1</td>
<td>0.01</td>
</tr>
<tr>
<td>Diborane</td>
<td>15</td>
<td>0.1</td>
<td>0.01</td>
</tr>
<tr>
<td>Dimethylamine</td>
<td>500</td>
<td>10</td>
<td>0.01</td>
</tr>
<tr>
<td>Fluorine</td>
<td>25</td>
<td>0.1</td>
<td>0.01</td>
</tr>
<tr>
<td>Formaldehyde</td>
<td>20</td>
<td>0.75</td>
<td>0.12</td>
</tr>
<tr>
<td>Hydrogen chloride</td>
<td>50</td>
<td>5</td>
<td>0.02</td>
</tr>
<tr>
<td>Hydrogen cyanide</td>
<td>50</td>
<td>10</td>
<td>0.02</td>
</tr>
<tr>
<td>Hydrogen fluoride</td>
<td>30</td>
<td>3</td>
<td>0.02</td>
</tr>
<tr>
<td>Hydrogen sulfide</td>
<td>100</td>
<td>20</td>
<td>0.08</td>
</tr>
<tr>
<td>Hydrazine</td>
<td>50</td>
<td>1</td>
<td>0.01</td>
</tr>
<tr>
<td>Methylamine</td>
<td>100</td>
<td>10</td>
<td>0.01</td>
</tr>
<tr>
<td>Methyl hydrazine</td>
<td>20</td>
<td>0.2</td>
<td>0.01</td>
</tr>
<tr>
<td>Nitric acid</td>
<td>25</td>
<td>2</td>
<td>0.02</td>
</tr>
<tr>
<td>Nitrogen dioxide</td>
<td>20</td>
<td>5</td>
<td>0.03</td>
</tr>
<tr>
<td>Phosphine</td>
<td>2</td>
<td>0.1</td>
<td>0.01</td>
</tr>
<tr>
<td>Phosphine</td>
<td>50</td>
<td>0.3</td>
<td>0.01</td>
</tr>
<tr>
<td>Sulfur dioxide</td>
<td>100</td>
<td>5</td>
<td>0.06</td>
</tr>
<tr>
<td>Trimethylamine</td>
<td>200</td>
<td>10</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Fig. 35 Color difference maps of 20 representative TICs at their IDLH. Reproduced with permission from ref. 204.

Fig. 36 HCA dendrogram for 20 TICs at IDLH concentrations and a control. All experiments were performed in septuplicate; no confusions or errors in clustering were observed in 147 trials. Reproduced with permission from ref. 197.
Fig. 37 Limits of recognition for TIC identification. Left: the effect of concentration on array response to NH₃, SO₂, and HCN. Right: HCA of a subset of TICs demonstrates a limit of recognition well below 5% of the PEL. Reproduced with permission from ref. 204.

Several reports of more specialized arrays for specific subsets of TICs have also been published. For example, Sen et al.²³⁵ developed a disposable colorimetric sensor array which can detect H₂S concentrations in the range of 50 ppb to 50 ppm at ambient temperature. Bang et al.²⁰⁷ prepared nanoporous silica microspheres incorporating chemoresponsive dyes and used an array of these to detect and quantify ammonia gas at its IDLH (immediately dangerous to life or health), PEL (permissible exposure limits), and 0.1 PEL concentrations with a reported (immediately dangerous to life or health), PEL (permissible exposure limits), and 0.1 PEL concentrations with a reported IDLH (Fig. 35 and 36). Limits of detection limits (listed in Table 1) were generally well below the PEL (in most cases below 5% of PEL) and are typically in the low ppb range. The colorimetric sensor array was not responsive to changes in humidity or temperature over a substantial range. The array performed well in the presence of various common potential interferents and has shown excellent stability and reproducibility.

While LODs are defined absolutely with respect to S/N, that only defines when one can determine that some analyte is present. Limits of recognition are much more important, but they are also library dependent. Fig. 37 demonstrates a limit of recognition for a subset of TICs is well below 5% of their PEL, which becomes of interest for epidemiological studies. Consistent with the array’s ability to discriminate among many possible TICs over many possible concentrations, PCA and LDA confirmed the high dimensionality of the colorimetric sensor array with 17 PCA dimensions required to capture 95% of the variance.¹⁹⁷

4.3 Explosives detection

Triacetone triperoxide (TATP), one of the most dangerous primary explosives, has emerged as an explosive of choice for terrorists in recent years. TATP is easily produced by an acid catalyzed reaction of acetone with hydrogen peroxide.²³⁸ Owing to lack of UV absorbance, fluorescence or facile ionization, TATP is difficult to detect directly.²³⁹ Techniques that are able to detect generally require expensive instrumentation, need extensive sample preparation, or cannot detect TATP in the gas phase. Lin and Suslick¹⁸¹ reported a new method for determination of TATP vapor using a colorimetric sensor array. In this method, the gas stream containing TATP vapor is decomposed by a solid acid catalyst (Amberlyst 15) and the resulting H₂O₂ vapor, being kinetically much more reactive, is easily detected by redox indicators (Fig. 38).

TATP was detectable even at very low concentrations using this technique, with an LOD below 2 ppb (i.e., <0.02% of its saturation vapor pressure). Common potential interferences (e.g., humidity, personal hygiene products, perfume, laundry supplies, volatile organic compounds, etc.) did not generate an array response, and the array could also differentiate TATP from other chemical oxidants (e.g., hydrogen peroxide, bleach, t-butylhydroperoxide, peracetic acid) (Fig. 26).

Kostesha, Alstrom et al.²⁴⁰,²⁴¹ have reported very briefly on a colorimetric sensor array capable of detection of explosives’ vapors, but no details were given on dyes used in their array. In this work, classification ability of K-nearest neighbor (KNN), artificial neural networks (ANN) and sparse logistic regression (SLR) methods were compared.

4.4 Aqueous analytes

If a colorimetric sensor array is printed on a hydrophobic membrane and the dye formulations sufficiently hydrophobic, then upon immersion into an aqueous solution containing organic compounds, the sensor array will respond to the volatile vapors of solutes. Zhang and Suslick²⁹⁹ prepared a simple colorimetric sensor...
array that was able to probe different organic compounds at very low concentrations (below 1 μM) in water. As shown in Fig. 39, unique fingerprints were observed for a wide range of dissolved organic compounds.

Monitoring toxic metal ions in water has also been accomplished using a nanoporous pigment array. Feng et al.\textsuperscript{151} reported a CSA for identification trace heavy metal ions (including Hg, Pb, Cd, Zn, Ag, As, Ni and Cu) at waste water-discharge standard concentrations. Suitable chemoresponsive dyes were immobilized using ormosil formulations; in this case, five separate probes were analyzed asynchronously through a filtration method and organized into an array. Discrimination of heavy metal ions was performed without interference of Na\textsuperscript{+}, K\textsuperscript{+}, Ca\textsuperscript{2+}, Mg\textsuperscript{2+} ions and good repeatability and high stability was obtained for this sensor.

Recognition of amino acids in aqueous solutions has also been accomplished using a colorimetric sensor array. Dan-Qun et al.\textsuperscript{242} used a 6 × 6 array to distinguish among amino acids based on differences between their chemical properties and specific residue structures. Ten natural amino acids including glycine (Gly), valine (Val), methionine (Met), proline (Pro), serine (Ser), tyrosine (Tyr), glutamine (Gln), glutamate (Glu), lysine (Lys), and histidine (His) were identified within 5 min of exposure at concentrations of 375 μM (Fig. 40).

Fluorescent displacement assays have also been used for identification of amino acids in solution. For example, dansyl-modified β-cyclodextrins bearing a metal binding site, with pendant ε-amino acids, were employed for selective identification of unmodified amino acids in aqueous buffer at neutral pH\textsuperscript{243,244} and as enantioselective fluorescence sensors for the discrimination of enantiomers of the amino acids valine and proline. The best conditions to perform enantiomeric analyses I used fluorescence quenching by the copper(II)-amino acid complexes in a fluorescence microplate reader.

Fluorometric displacement assays using complexometric solution sensor arrays have been used for identification of metal ions. In 2003, Mayr et al.\textsuperscript{152} developed an eight-component fluorometric array for discrimination among five separate cations using cross-reactive complexometric fluorophores in solutions held in a microtiter plate. The wells were monitored simultaneously using a CCD camera and a set of fiber-optic cables to reduce the imaged area (Fig. 41). The fluorescence decay profile of the indicator is referenced against the phosphorescence of an added inert reference dye at a different wavelength. The source of illumination were blue LEDs, one for each well. The assembly allows the detection of dye concentrations in the nanomoles-per-liter range without amplification and the acquisition of 96 wells simultaneously. Some discriminatory ability was shown by this straightforward liquid-phase array, but limited statistical analysis was performed; classification accuracy of 82 to 93% was reported. A gray-scale image showing the raw

Fig. 39  Color change profiles with the base-sensitive sensor array for representative aqueous solutions of organic compounds (all amines 10 mM, all others 50 mM, in pH 7 phosphate buffer). Reproduced with permission from ref. 199.

Fig. 40  Average color change profiles of 10 amino acids visualized as color difference maps. Reproduced with permission from ref. 242.

Fig. 41  (a) Fiber-optic adapter for fluorescent imaging of 96 well microtiter plates; (b) schematic of the apparatus showing LED array illumination and CCD camera detection. Reproduced with permission from ref. 152.

Fig. 42  Greyscale output of eight-component array after exposure to various concentrations of metal ions. Analytes vary by column, and included some multi-ion solutions; specific concentrations are given in ref. 152. Of note, columns 5–9 are set at 1 μM Ca\textsuperscript{2+}, Cu\textsuperscript{2+}, Ni\textsuperscript{2+}, Zn\textsuperscript{2+}, and Cd\textsuperscript{2+} respectively. Reproduced with permission from ref. 152.
fluorescence response to several analyte sets are shown as Fig. 42.

4.5 Applications to complex mixtures

4.5.1 Foods and beverages

Colorimetric sensor arrays. The analysis of complex mixtures presents a difficult challenge even for the most sophisticated analytical techniques, and the ability to discriminate among closely similar such mixtures often remains problematic. Foods and beverages are characteristic in the complexity of their compositions. A component by component analysis is generally unwanted and often impractically difficult, given the hundreds of different compounds found in edible materials. Instead, one is more interested in questions of authenticity, contamination, and food processing quality control. For these goals, the sort of fingerprinting that sensor arrays provides can prove extremely valuable.

Coffee provides a readily available archetype of such highly multicomponent systems. While unroasted, green coffee contains more than 300 volatile compounds, while more than 1000 volatile compounds have been identified for roasted coffee including carboxylic acids, alcohols, aldehydes, alkanes, alkenes, aromatics, esters, furans, ketones, lactones, oxazoles, phenols, pyridines, pyrazines, pyroles, thiazoles, and thiophenes.245–247 Furthermore, the roasting of coffee beans is highly dynamic, and the processes that develop the flavor and aroma of coffee are strongly time and temperature dependent.

Suslick, Feng and Suslick248 made use of the same sensor array developed for TIC identification to the analysis of coffee aromas. The color changes of the sensor array were used as a digital representation of the array response and analyzed with standard statistical methods. PCA revealed that the sensor array has exceptionally high dimensionality with 18 dimensions required to define 90% of the total variance and 25 dimensions for 95%. In quintuplicate runs of ten commercial coffees and controls, no confusions or errors in classification by HCA were observed in 55 trials (Fig. 43). In addition, the effects of temperature and time in the roasting of green coffee beans were readily observed and distinguishable with a resolution better than 10 °C and 5 min, respectively.

Zhang and Suslick had earlier applied a very similar approach to commercially available soft drinks using an early version of their array containing only 25 chemically responsive dyes printed on a hydrophobic membrane.249 Fourteen commercial soft drinks were analyzed and facile identification of all of the soft drinks was readily achieved using comparison of the color change profiles or a PCA score plot (Fig. 44) Using a HCA dendrogram, the misclassification rate was <2%, and even very similar sodas were easily differentiated. In addition, the monitoring of soft drinks as they degas or upon dilution also proved to be possible. This work demonstrated the potential of colorimetric sensor array technology for quality control.
assurance/control applications of sodas and perhaps other beverages as well.

In a similar application, Zhang et al. showed that colorimetric sensor arrays were able to differentiate among 18 brands of beer. Differentiation between ales and lagers (Fig. 45) was without error, and even among very similar beers proved straightforward with an error rate of identification <3%. In addition, differentiation of pristine beer from the effects of watering or de-carbonation proved possible. These results suggest that colorimetric sensor arrays may prove useful for QA/QC applications.

Ya et al. also analyzed five commercial baijiu (a Chinese distilled alcoholic beverage) using a simple colorimetric artificial nose. The presence of chemoresponsive dyes containing porphyrins and porphyrin derivatives provided a unique pattern of color changes in response to baijiu. With the aid of PCA and HCA, classification of types of baijiu was done according to their trace components. Using linear discriminant analysis (LDA), flavor styles were determined with 100% accuracy.

Y. T. Chang and co-workers have developed spectrophotometric microplate assays for evaluation of tap and bottled water using 45 dyes including pH indicators, metal ion complexometric indicators, and substituted quinone dyes known to be responsive to various organic and inorganic species. By monitoring the optical spectra of the microwell plates, they were able to classify water samples in terms of their place of origin, metal ion content, and carbohydrate content.

Detection and quantification of sugars and sweeteners is very important in real-time food quality. Musto, Lim, and Suslick have developed colorimetric sensor arrays for detection and quantification of sugars and artificial sweeteners by immobilizing suitable chemoresponsive dyes in insoluble nanoporous pigments. The sensor array was able to accurately determine 14 sugars and sweeteners at millimolar concentrations at pH 7.4 (Fig. 46). The concentrations of sugars and sweeteners could also be determined over at least a five-fold range, and glucose concentrations were measurable over the full range of clinical importance for blood sugar determinations. The sensor array worked as well in tea infusions. Ghosh et al. applied the same concepts later to a microplate liquid ensemble of boronic acid derivatives and pH indicators with similar success and were able to show quantitative analysis of sugar concentrations.

Monitoring of food freshness, especially of meats, poultry and fish, are an obvious application of sensor arrays. Huang et al. constructed a colorimetric sensor array by printing the nine chemoresponsive dyes on a reverse phase silica gel plate. These dyes were sensitive to volatile compounds produced during spoilage of fish. PCA and neural network techniques were applied for classification of the degree of spoilage, allowing evaluation of the fish freshness with the accuracy of 87.5%. Recently Salinas et al. reported an optoelectronic nose constructed of 16 pigments able to identify the age of chicken meat.

**Indicator displacement assays.** Anslyn and co-workers have developed a general displacement strategy using serum-albumin as a differentially selective receptor capable of binding both fluorescent reporters and a variety of non-polar analytes (e.g., fatty acids). The resultant probe was able to differentiate among fatty acids (palmitic acid, oleic acid, stearic acid, and linoleic...
acid) and among edible oils (sunflower oil, hazelnut oil, canola oil, extra virgin olive oil, and peanut oil). A graphical representation of their data using PCA is shown as Fig. 47, showing clear distinction among analyte species. It is worth noting that the PCA is very strongly dominated by a single component (~90% of total variance), and that dimension essentially represents the hydrophobicity of the analytes.

In a similar concept, Anslyn and coworkers used a set of metallo-histidine peptides in order to discriminate among polyphenol-based wine flavonoids. In this indicator displacement assay, the probes were based on replacement of a catechol dye from a Cu²⁺ center bound to the peptide; the design strategy (shown as Fig. 48) is quite straightforward and specific due to significant structural similarity between the indicator dye and chosen analytes. The resulting discrimination, however, was only moderate.

Much more impressive results were obtained by the same group using a diversified set of colorimetric indicators and synthetic boronic acid and guanidinium functionalized receptors. The authors showed the ability of the array to discriminate malate, tartrate and citrate and then using spectrophotometric data at three wavelengths (the indicator maxima), they were able to discriminate accurately among six different red wines, as shown in Fig. 49. In contrast with the earlier paper, these more chemically diverse displacement assays provide a much higher dimensionality of data (three dimensions include 88% of the discriminatory power) and a much improved distinction between the wines.

4.5.2 Proteins. There have been several recent studies reported on the optical sensor array identification of single proteins in solution. While it is not clear what application single protein identification may have, in some cases the approaches are intellectually clever. One may hope that such arrays might prove useful for detecting a specific rogue protein against a constant complex mixture as found in plasma; cross-reactive sensor arrays, however, are unlikely to be able to achieve such a goal at biomedically relevant concentrations. Such tests remain to be reported.

For colorimetric sensor arrays, Hou et al. tested an array modeled closely on Suslick’s work using a series of porphyrins and indicator dyes capable of rapid interaction with proteins.
The array produced distinct patterns in response to each protein which permitted accurate identification of the pure and mixed proteins as well as denatured proteins. Fig. 50 shows two dimensional PCA plots which show excellent discrimination among the proteins, especially given that only ~ 64 to 82% of the total variance is captured in a two dimensional plot. Remarkably, one is able to discriminate thermally denatured proteins from the native form rather easily.

In another example of a pseudo-array data analysis using sequential array sensing of parallel solutions, displacement or differential assays have also been used to identify protein solutions. The Rotello group generated multiple differential sensing parallel arrays using gold nanoparticles with the aim of discriminating among a variety of biological samples; this work has been both succinctly and exhaustively reviewed. Several different fluorophores have been used. The interaction of analytes with nanoparticles adds additional unique characteristics, as shown schematically in Fig. 51.

For detection of seven common proteins (bovine serum albumin, cytochrome c, alkaline phosphatase, acid phosphatase, subtilisin A, lipase, and β-galactosidase) Rotello et al. prepared a sensor array based on six cationic functionalized gold nanoparticles (AuNPs) and an anionic PPE (poly(p-phenyleneethynylene)) polymer. The fluorescence of PPE polymer was quenched by interactions of AuNPs with polymer, but addition to a protein solution displaced the polymer and enhanced the observed fluorescence. By using different functionalization, one could alter the protein affinity to the surface of nanoparticles and, thus, producing distinct fluorescence response patterns. Fluorescence responses to individual proteins were fingerprinted, giving a detection limit of 4 to 215 nM depending on protein size.

Of particular note, Rotello used a gold nanoparticle-fluorophore system made from gold nanoparticles conjugated to green fluorescent protein (GFP-NP) to act as the displaced indicator. Using a protein as the displaced indicator is a straightforward method of discriminating among peptide-based analytes due to the molecular similarity between the incoming analyte and the displaced probe. This method was used for discrimination among several types of human serum proteins (Fig. 52), and then in a subsequent publication to mammalian cells as discussed in the last section.

Li et al. used a set of functionalized Fe nanoparticles to discriminate among several proteins. This work used only two
probes for discrimination, leading to an inherently low maximum dimensionality. In practice, the discrimination ability of the probes comes almost exclusively from a single dimension, as shown in Fig. 53. The authors state that surface charge, protein size, and surface hydrophobicity all play roles in discrimination of these species; they may all contribute in a composite sense to the principal component, but the one-dimensional nature of the data establishes that this is not truly an array sensor. The origin of the specificity is not clear; proteins with very similar pI values are differentiated, and there is no correlation with protein size or surface charge. A more detailed data about surface hydrophobicity may provide some further insight.

### 4.5.3 Intra-cellular sensing using ratiometric fluorometry.

In a sense, the use of two or more fluorophores in a single probe is also array sensing. So ratiometric fluorescent probes, where one has two fluorophores excited at the same wavelength, but emitting at different wavelengths, permits an array-like response that is self-calibrating. Let us examine a few recent examples of intra-cellular sensing using ratiometric fluorometry.

Intra-cellular temperature probes provide an interesting example. Small local temperature changes may have significant effects on the cellular signaling pathways and their nanoparticle uptakes; thus, the development of novel temperature nanosensors may prove very useful for intra-cellular profiling and imaging. Peng et al. prepared a ratiometric type of fluorescent nanoparticles with a green and a red fluorescence from a single wavelength excitation; these nanoparticles showed the capability of monitoring temperature variations in the physiological temperature range (i.e. 25–45°C), as illustrated in Fig. 54.

![Fig. 54](image-url)

**Fig. 54** (a) Schematic representation of temperature-sensing nanoparticles containing a random dispersion of Eu-DT (structure shown as b) and OASN (structure shown as c) in a silylated poly(methyl methacrylate) matrix (BTD-PMMA), which is covered with a silica outer shell; (d) ratiometric fluorescence response; (e) temperature calibration plot. Reproduced with permission from ref. 278.

Monitoring of the pH inside live cells is also of importance for probing living cell functions. Peng et al. employed ratiometric fluorescent nanogels for monitoring pH variations (Fig. 55). The hydrogel was prepared from polyurethane containing both hydrophilic and hydrophobic domains, thus the nanogel chains rearranged to form a three-dimensionally stable nanostructure when placed in water. It is noteworthy to mention that the volume of the nanogels were not affected by pH, so the efficiency of FRET was constant during in vitro pH sensing. The developed pH probe is suitable for monitoring the pH in the range of 6–8. Fig. 55b shows the fluorescence emission spectra of the nanogels at various pH values upon excitation at 450 nm. Calculation of the ratio of the observed emission intensities of the dyes in various pH values is presented in Fig. 55, showing around nine fold variation in the ratio on going from pH 5 to pH 9.

A final example simultaneously measured oxygen content and pH value in bacterial cultures (Pseudomonas putida), Kocincová et al. prepared and employed a “dual” sensor based on organosilica microparticles (for oxygen sensing) and polymethacrylate derivative embedded into a polyurethane hydrogel (for pH sensing).

### 4.5.4 Detection and identification of bacteria.

The detection and identification of bacteria are pressing problems in both medicine and industry. Bacterial infections are involved in food poisoning, hospital-acquired infections, and other areas that are of great concern for public health, and sepsis remains one of the leading causes of death even among first world nations. In industry, many products must be screened after manufacture for bacterial contamination before they may be released and as a consequence regulation of the food industry must be particularly stringent. Existing methods
for identification of pathogenic bacteria are limited by the necessity of long culturing times, the need for highly trained laboratory personnel, or the requirement of expensive and high-maintenance equipment.289–294

Bacteria stink: that is, they produce volatile organic compounds (VOCs) to which the mammalian olfactory system is highly responsive. Consequently, an experienced microbiologist can readily identify many bacteria by smell. Applications of prior electronic nose technology, however, have been limited by the low dimensionality of traditional sensor arrays (e.g., metal oxides) and have achieved only modest success, even when attempting to classify small numbers of bacterial species.295–297

Using a disposable colorimetric sensor array, Carey et al.298 were able to identify different species and specific strains of human pathogenic bacteria based on volatile compounds produced during bacterial growth. In 164 trials monitoring bacterial growth as a function of time, they were able to identify 10 strains of bacteria including Enterococcus faecalis and Staphylococcus aureus and their antibiotic-resistant strains during 10 hours of culture time with 98.8% accuracy as evaluated by jackknifed LDA using time-stacked data. Fig. 56 shows the color difference maps and time evolution of the array response to the head-gas mixture produced in sealed Petri dishes. Fig. 57 shows a PCA score plot using the three most important dimensions, which account for only 79% of the total variance.

Rotello and co-workers used an aqueous solution based displacement assay to identify bacteria in liquid growth media using the fluorescence quenching of fluorophores attached to gold nanoparticles.299 They made use of cationic gold nanoparticles with conjugated polymer fluorophores as their differential fluorescence probe, whose response is dictated by the binding strength of the bacterium to the gold nanoparticle. By manipulating the surface chemistry of gold nanoparticles and the constitution of the conjugated polymer, they generated array-like data (albeit one solution at a time). Nine bacterial species and 3 strains of E. coli were examined in 64 trials with a 95% accuracy of classification by LDA (Fig. 58).

The immobilized fluorescent bead strategy that Walt developed for use with microwelled optical fiber bundles (discussed earlier) has also been recently applied to bacterial identification. Fixed arrays of this sort are difficult to reuse

Fig. 56 (a) Color difference maps and time response profile resulting from colorimetric sensor array exposure to a growing culture of E. coli, American Type Culture Collection (ATCC) #25922. (b) The color change values versus time plotted for all color channels ($D_R$, $D_G$, and $D_B$ values for each spot, i.e., 108 color channels) at each time point. (c) Color difference maps for 10 different bacterial strains resulting from colorimetric sensor array exposure to Petri dish growing cultures after 480 min. Reproduced with permission from ref. 298.

Fig. 57 PCA score plot using the three most important principal components based on 164 trials of 10 bacterial strains and controls. The resolution between bacterial classes is in fact much better than can be shown by any three-dimensional PCA plot because the first three principal components account for only 79% of the total variance. ● S. aureus; ● MRSA; ● S. epidermidis; ● S. scuri; ● P. aeruginosa; ● E. faecium; ● E. faecalis; ● E. faecalis VRE; ● E. coli 25922; ● E. coli 53502; ● control. Reproduced with permission from ref. 298.

Fig. 58 LDA score plot for the fluorescence response patterns to bacteria as determined from solution displacement assays using gold nanoparticle displacement assays. The first two dimensions contain 96.2% of the variance. Reproduced with permission from ref. 299.
multiple times, in part due to photobleaching. Walt and coworkers have attempted to address this problem by using a free bead strategy,\textsuperscript{300} using optical trapping, the arrays are re-created between each use. Since the methods of signal transduction are identical to those used in the immobilized bead strategy, it is expected that performance will be similar. A false-color image showing raw data output using this method is shown as Fig. 59 for detection of three strains of \textit{E. coli}.

4.5.5 Cancer and disease diagnosis. Optical sensor arrays also have begun to find applications to medical diagnosis of disease. Different cell line produce different volatile metabolites, as discussed early with bacteria; this applies to any rapidly growing cells, and in principle the breath is in equilibrium with the volatiles produced within the body. Breath analysis has a long history as an underutilized diagnostic tool;\textsuperscript{301–303} limitations in analytical tools that are sufficiently sensitive, specific, or inexpensive have limited this approach in clinical or hospital applications. Electronic noses\textsuperscript{301,303,304} have certainly been examined for breath analysis, especially for diagnosis of lung cancer and of respiratory infections with some limited success. Colorimetric sensor arrays of the type developed in Suslick’s lab for bacterial identification\textsuperscript{298} too, have had some preliminary clinical success for breath diagnosis. Point of care diagnosis of bacterial sinusitis, as one example, has achieved classification accuracy as high as 90% in initial studies at the University of Pennsylvania medical school.\textsuperscript{305} Lung cancer screening using breath analysis has also been reported by Mazzone et al.\textsuperscript{304} at the Cleveland Clinic with promising initial results. In a study with 229 subjects (92 with lung cancer), individuals with different histologies could be accurately distinguished from one another (86.4% for adenocarcinoma vs. squamous cell carcinoma), and the accuracy of breath biosignatures could be optimized by incorporating clinical risk factors.\textsuperscript{304}

In an entirely different approach to cell differentiation, it is not surprising that different cell lines interact differently with different nanoparticles\textsuperscript{306–308} and that those interactions are strongly affected by the chemical nature of the nanoparticle surfaces, particularly as biomolecules adsorb onto nanoparticle surfaces forming a “protein corona”.\textsuperscript{309–311}

Taking advantage of nanoparticle-whole cell interaction, the parallel soluble fluorescent displacement assays that Rotello et al.\textsuperscript{272,312,313} have developed for protein identification also have been applied to differentiation between normal and cancerous whole cells. For example, green fluorescent protein coated gold nanoparticles array were used to differentiate types of mammalian cancer cells (HeLa (cervical), HepG2 (liver), MCF-7 (breast), and NT2 (testis)).\textsuperscript{312} The nanoparticle surface ligand structures for mammalian cell discrimination are shown as Fig. 60. A score plot summarizing the array’s discrimination ability\textsuperscript{272} is shown as Fig. 61.

---

**Fig. 59** False color images of raw three-component microbead array responses to three strains of \textit{E. coli}. Each column is one strain. Top row (a–d) shows the identity of the beads from their fluorescence, while bottom row (e–h) shows intensity response to the exposed analyte as a gray scale. The array is \( \sim 25 \times 25 \) \( \mu \text{m} \). Modified from ref. 300.

**Fig. 60** The six nanoparticle ligand structures used for mammalian cell discrimination. Differential sensing ability derives from displacement of GFP from the nanoparticles to release GFP in the presence of species with differential relative affinities for the gold nanoparticle with various surface ligands (NP1 through NP6). Reproduced with permission from ref. 272.

**Fig. 61** LDA score plot showing clear differentiation between healthy and metastatic mammalian cells and among cells from various organs. The blue/red shading is meant to delineate healthy/normal and metastatic/tumor cells. Reproduced from ref. 272.
5 Limitations, opportunities, and future challenges

The primary limitation of a sensor array, including the mammalian nose, is that it gives a composite response to complex mixtures. One does not get a component by component analysis from cross-reactive sensor arrays. Most chemists naturally assume that the goal with any complex mixture is to achieve a complete quantitative analysis of each component. But in fact, for multi-component analytes, there are generally multiple analytical goals: one often wants to know is this mixture the same as that one, is there one (or a few) component(s) that have changed against a constant complex mixture, is this mixture genuine or counterfeit, was this material processed correctly or not?

The primary strength of a sensor array, including the mammalian nose, is that it gives a composite response to complex mixtures. The ability to fingerprint complex mixtures is often greatly simplified by such a composite response. With recent advances in optical sensor arrays, exquisite fingerprinting of extremely similar mixtures or single analytes has been achieved over a wide range of analyte types.

Another characteristic of sensor arrays that probes chemical properties of analytes, rather than physical properties, is non-uniform intrinsic response to analytes. This too is a good news–bad news story. The good news is that most analytes of concern (e.g., toxic gases) are, essentially by definition, highly reactive and therefore easily detected, even at sub-ppm concentrations. This eliminates many of the problems of false-positives associated with traditional electronic nose and solid state chemical sensors. The bad news is that some analyses are interested in less reactive analytes. One solution, developed only recently, is to pre-react the analyte stream to produce, for example, partial oxidation products that are more reactive and therefore more easily analyzed by chemoresponsive sensors. More selective methods of “activating” these analytes, especially in the presence of a high concentration of interferents, would add significantly to the capabilities of current optical sensor arrays for gas analysis.

The alternative of incorporating even more promiscuous optical sensors has some merit as well; the difficulty is that the weak interactions usually used in monitoring such properties (e.g., physisorption onto surfaces, absorption into polymers) generally lead to lower sensitivity (e.g., solvatochromic shifts of polymer adsorbents are typically undetectable below ~0.1% of the saturation vapor pressure of an analyte).

For heterogeneous optical arrays designed for liquid sensing, contamination of analyte solutions due to probe solubility is a major problem. One is caught in a dilemma for solid state optical arrays that are to be immersed in liquids: the solvates must have access to the sensor medium, but the sensor dyes must not dissolve into the liquid. In principle, one may overcome this problem by immobilization of dyes into sol-gel formulations, but in practice immobilization is sometimes imperfect.

Solution phase array sensing is carried out by a parallel analysis of multiple aliquots of the analyte solution, each with an added, different homogeneous probe, e.g. using microwell plates. This is a cumbersome procedure, and one that should yield to improved technology using heterogenization of the soluble probes, including covalent attachment to the probe substrate.

All in all, however, one can only be impressed by the development of optical sensor arrays over the past decade. Colorimetric sensor arrays have demonstrated excellent potential for complex systems analysis in real-world applications and provide a novel method for discrimination among closely similar complex mixtures.

The commercialization of optoelectronic nose technology is underway in several companies and one expects that significant markets are likely to develop in the future.

Acknowledgements

This work was supported through the NIH Genes, Environment and Health Initiative (U01ES016011), the Department of Defence (Contract N41756-12-R-4767), and the NSF (CHE 1152232).

Notes and references

63 C. E. W. Hahn, Analyst, 1998, 123, 57R–86R.
113 I. M. Kolthoff, Acid Base Indicators, Macmillan, New York, 1937.


Review Article


