

Article

# Genetic Algorithm Design of MOF-based Gas Sensor Arrays for CO<sub>2</sub>-in-Air Sensing

Brian A. Day <sup>1</sup> and Christopher E. Wilmer <sup>1,2,\*</sup> 

<sup>1</sup> Department of Chemical and Petroleum Engineering, University of Pittsburgh, 3700 O'Hara St, Pittsburgh, PA 15261, USA; brd84@pitt.edu

<sup>2</sup> Department of Electrical and Computer Engineering, University of Pittsburgh, 3700 O'Hara St, Pittsburgh, PA 15261, USA

\* Correspondence: wilmer@pitt.edu

Received: 19 December 2019; Accepted: 29 January 2020; Published: 10 February 2020



**Abstract:** Gas sensor arrays, also known as electronic noses, leverage a diverse set of materials to identify the components of complex gas mixtures. Metal-organic frameworks (MOFs) have emerged as promising materials for electronic noses due to their high-surface areas and chemical as well as structural tunability. Using our recently reported genetic algorithm design approach, we examined a set of 50 MOFs and searched through over  $1.125 \times 10^{15}$  unique array combinations to identify optimal arrays for the detection of CO<sub>2</sub> in air. We found that despite individual MOFs having lower selectivity for O<sub>2</sub> or N<sub>2</sub> relative to CO<sub>2</sub>, intelligently selecting the right combinations of MOFs enables accurate prediction of the concentrations of all components in the mixture (i.e., CO<sub>2</sub>, O<sub>2</sub>, N<sub>2</sub>). We also analyzed the physical properties of the elements in the arrays to develop an intuition for improving array design. Notably, we found that an array whose MOFs have diversity in their volumetric surface areas has improved sensing. Consistent with this observation, we found that the best arrays consistently had greater structural diversity (e.g., pore sizes, void fractions, and surface areas) than the worst arrays.

**Keywords:** metal-organic framework; computational design; carbon dioxide; gas sensing

## 1. Introduction

There has been continued interest in addressing the challenge of recreating the gas-sensing capabilities of a biological nose via a gas-sensing array, known as an electronic nose [1–3]. Unlike other gas-sensing technologies, such as gas chromatography (GC) or mass spectroscopy (MS), which can be slow, expensive, and are largely not portable, an electronic nose can be small, cheap, and provide real-time sensing [1,4]. However, designing an electronic nose that can rival a biological nose is enormously difficult, in part because gas mixtures in ambient environments are highly complex, often including hundreds of species that competitively bind to the same receptors [3]. Nevertheless, there are numerous applications where an electronic nose would not need to compete with a biological nose, but rather perform in ways where current gas-sensing technologies are insufficient, whether due to a long response time or lack of sensitivity, such as in disease detection [4–7], or due to a lack of portability, such as for emergency response [8], agriculture [6,9,10], and mobile robots [10–12].

One application that is seeing renewed interest is in the detection of carbon dioxide, primarily because of its role in climate change [13–15]. There is, consequently, an obvious motivation to be able to reliably detect and monitor levels of CO<sub>2</sub> in the atmosphere. Additionally, with geologic carbon sequestration gaining traction as a viable carbon capture and storage (CCS) technology to mitigate the greenhouse gas effect of CO<sub>2</sub>, the need to detect leaks from storage sites grows with it [16–18]. Other applications include the detection of CO<sub>2</sub> in confined spaces where it can rapidly accumulate,

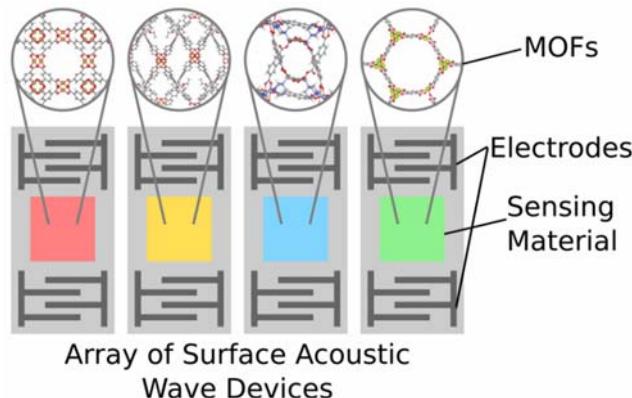
especially when tanks of CO<sub>2</sub> or dry ice are being used [19]. And even in areas where CO<sub>2</sub> is less likely to accumulate rapidly, there are concerns over its role as an indoor pollutant [8,20,21].

Exposure to elevated levels of CO<sub>2</sub> in air poses a two-fold threat, acting as both an asphyxiant by displacing oxygen, and as a toxicant, both with potentially deadly consequences. At levels greater than 5%, it can result in the development of hypercapnia, a build-up of CO<sub>2</sub> in the bloodstream, and respiratory acidosis, an inability to clear excess CO<sub>2</sub> from the lungs [8,20,22]. At concentrations of 10% and greater, exposure can result in convulsions, coma, and even death. And at levels of 30% and greater, exposure can lead to a loss of consciousness in only a matter of seconds [8,20]. Clearly, rapid, sensitive, and portable (even wearable) CO<sub>2</sub> sensors would benefit many people.

Recently, a number of studies have investigated the use of metal-organic frameworks (MOFs) for improving electronic noses [23–27]. MOFs are promising materials for gas adsorption applications due to their nanoporous nature, high internal surface areas, and ease of tunability [28–34]. Moreover, as crystalline materials, they lend themselves conveniently to accurate computational modeling not possible with their porous amorphous counterparts [35–37].

Practically, however, there are still many significant obstacles to overcome in developing MOF-based electronic noses. While tens of thousands of MOFs have been reported in the literature, only a small fraction have simultaneously been amenable to thin-film deposition and also possess the stability required for a practical device. However, even if those material-property obstacles were overcome, an important design challenge has remained largely overlooked: what combination of MOFs optimizes performance? Although it is intuitive that adding more sensing elements to an array should improve performance, to what degree the performance can be improved has not been widely explored [38]. Similarly, there has been little work on systematically finding the best combinations of MOFs for an array, and with thousands of MOFs to choose from, and thus well over  $10^{50}$  possible arrays, determining the top performing arrays is highly non-trivial. A recent paper by Sturluson et al. made significant progress in this area by computationally designing MOF-based gas sensing arrays for the detection of CO<sub>2</sub> and SO<sub>2</sub> in dilute conditions [27]. In their paper, they calculated the Henry's coefficients, which are used to predict adsorption of gases in dilute concentrations, for each gas/MOF to design arrays with maximal complementarity, and consequently the best signal.

In this paper, we use a computational methodology to find optimal MOF-based gas-sensing arrays for CO<sub>2</sub> detection in ambient air, where the MOFs are assumed to be deposited as thin films on sensors that detect changes in mass, such as surface acoustic wave (SAW) devices, as shown in Figure 1, or quartz crystal microbalances (QCMs). The number of possible arrays one could construct grows as  $N!$  where  $N$  is the number of MOFs to choose from in a given library (the size of the search space also depends on the maximum array size allowed, see the Supplementary Information Section 3.2 for more details). In our case, from a library of 50 MOFs where arrays could have any size from 1 through to 50 MOFs, one could construct over  $1.125 \times 10^{15}$  unique arrays. Since a brute force search clearly could not be used for such a large space, we used a genetic algorithm strategy to quickly identify high performing arrays of different sizes, while only exploring a fraction of the space. Genetic algorithm searches were constrained to fixed time (i.e., fixed number of arrays considered) and so the computational cost was independent of the size of the search space. With the resulting data, we examined the physical properties of the MOFs in each array to find which combinations of properties are typically found in the best (or worst) arrays for CO<sub>2</sub>-in-air sensing.



**Figure 1.** Sample four-element array of surface acoustic wave devices with MOFs as sensing materials.

## 2. Materials and Methods

To model  $\text{CO}_2$  in air, we studied a set of ternary gas mixtures containing  $\text{CO}_2$ ,  $\text{O}_2$ , and  $\text{N}_2$ . The compositions of  $\text{CO}_2$  and  $\text{O}_2$  ranged from 0% to 30%, and the composition of  $\text{N}_2$  ranged from 40% to 100%, each in increments of 1%. This resulted in 961 unique gas mixtures.

Using the molecular modeling software RASPA [39], we ran grand canonical Monte Carlo (GCMC) simulations to calculate the adsorption of each gas mixture in a set of 50 MOFs at a temperature of 298 K and a pressure of 1 bar, reflecting ambient conditions. We used the same set of 50 MOFs as was used in a prior work on methane-in-air sensing [24], which were selected from the CoRE MOF database [40] with the aim of having a diverse representation of surface areas and void fractions.

To model electrostatic interactions, which are important for accurately predicting  $\text{CO}_2$  and, to a lesser extent,  $\text{N}_2$  adsorption, we assigned partial charges to the atoms of the MOF frameworks via the EQeq method [41]. Similarly, the molecule parameters of the gases also included partial charges, and the forcefield which we used, TrAPPE, has been shown to accurately simulate these effects [42].

Once we had a complete set of data from the adsorption simulations, the next step was to design and evaluate gas sensing arrays with the elements, emulating SAW sensors that have had MOF thin films deposited on them. SAW sensors measure changes in the mass of these thin films due to gas adsorption. To recreate this type of signal, we used a subset of the simulated adsorption values as the sensor outputs here, those values coming from simulations of 5%  $\text{CO}_2$ , 20%  $\text{O}_2$ , and 75%  $\text{N}_2$ . By comparing the sensor output from each array element to our library of simulated data, we can assign a probability to each of the 961 possible gas mixtures.

Note that since sensor outputs are simply a subset of the simulated data, there is always a composition whose simulated adsorption values are a perfect match. However, since there is always some error in measurements from real sensor devices, we account for this by artificially imposing measurement error on the sensor output, which ensures that no gas mixture can be predicted with 100% certainty.

An overview of the process of assigning these probabilities is as follows: one MOF at a time, we take the output signal (i.e., mass) associated with that MOF and create a truncated normal probability curve centered about the mass, with standard deviation 5% of that mass. The intention of using a truncated probability distribution rather than a true normal distribution is to account for the fact that adsorption will always result in an increase in mass. Consequently, the lower bound is set at 0, and the upper bound is set far beyond the highest simulated mass present in the dataset.

For each composition, we assign a probability based on where simulated mass sits on the truncated probability curve. Since each composition is assigned a probability independent of each other, the sum of all probabilities does not necessarily equal 1. However, since the intention of this process is to determine which simulated composition we have exposed the array to, we normalize the assigned probabilities for each MOF so that now their sum equals 1. This process is repeated for each MOF until we have one normalized probability value for each composition for each MOF.

Now that we have the probabilities for each MOF, we need to determine the probabilities for arrays. Fortunately, this process is straightforward. For each composition, we simply multiply all of the normalized probabilities for each MOF with each other, resulting in a non-normalized array probability for each composition. As before, we normalize these probabilities so that they sum to 1. With this information, we can now say which of the simulated gas mixtures best matches the sensor output of any given array. For a more complete description about this process, please refer to the supplementary information, as well as to the previous work which this is building upon [24–26].

In order to quantify the array performance beyond simply looking at the probability of the gas mixture we are testing for, we used a metric known as the Kullbeck-Liebler Divergence (KLD) [43], which is calculated (in units of bits) as follows:

$$KLD = \sum_{i=1}^N P_i \cdot \log_2(P_i \cdot N) \quad (1)$$

where  $P_i$  is the probability of composition  $i$  and  $N$  is the total number of possible compositions. In essence, the KLD tells us how much better we predict a composition over random chance, with a higher KLD corresponding to better prediction.

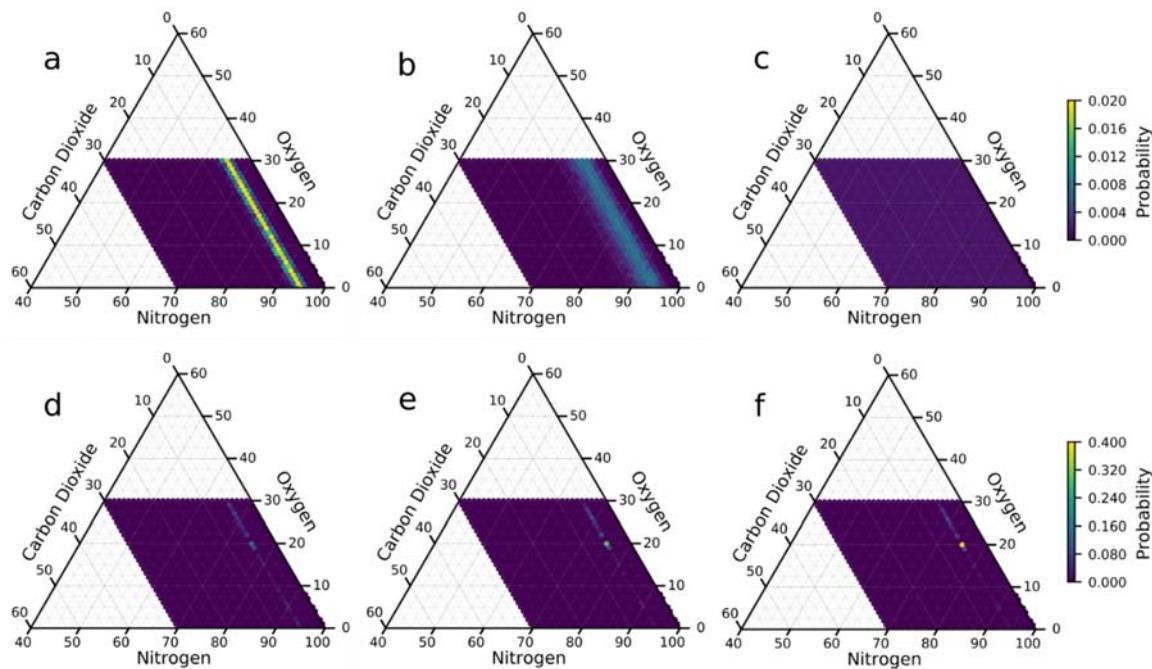
One critical advantage of the KLD is that it is calculated without any knowledge of the experimental gas mixture; that is to say, it is not an assessment of the accuracy of the prediction, but rather an assessment of the uniqueness of the signal. Furthermore, it is a metric that is transferrable to real sensor devices, not just the computational devices we examine in this paper. As such, the KLD score could even be used as a metric for determining the minimum performance requirements for a given array. Additional details on the KLD score are given in the supplementary information.

Lastly, for designing and screening arrays, we had two approaches available to us: a brute force approach, in which we evaluate all possible arrays of a given size, and a genetic algorithm approach [24], in which we continually update a set of arrays with a mutation strategy in order to seek the best performing arrays. Even with only 50 MOFs to choose from, there are over  $2.1 \times 10^6$  possible 5-element arrays. For 25-element arrays, the number of possibilities grows dramatically to over  $1.25 \times 10^{14}$  arrays; hence, the need for the genetic algorithm. Although many other search algorithms could also have been used, we chose to use a genetic algorithm for its relative simplicity, easy implementation, and broad familiarity in the scientific community. Additionally, since the objective function we chose for the genetic algorithm simply seeks to maximize the KLD, it makes no implicit assumptions about which features of an array are desirable and can thus be used to develop intuition for array design, as is done in this work. A detailed description of the genetic algorithm approach can also be found in the supplementary information.

### 3. Results

#### 3.1. Complementary Effect of Sensing Elements in a Single Array

As a baseline for understanding how well the MOFs we had selected perform for  $\text{CO}_2$  sensing, we examined the performance of all 1-element arrays, which is to say we examined how well each of the MOFs can individually predict the composition as a stand-alone sensor. Given the mixture of  $\text{CO}_2$ ,  $\text{O}_2$ , and  $\text{N}_2$ , an interesting result appears; all of the MOFs studied show a unique sensitivity to  $\text{CO}_2$  over the other two gases (see Figure 2a–c). This is likely because  $\text{CO}_2$  has a significant quadrupole moment and adsorbs more strongly than either  $\text{O}_2$  or  $\text{N}_2$ . Figure 2a–c shows, from all 50 of the 1-element arrays, the best performing MOF (based on its KLD score) was Mg-MOF-74 [44], the 25th best was  $\text{Cu}_4\text{I}_4(\text{DABCO})_2$  MOF [45], and the worst was the  $\text{La}(\text{PODC})_{1.5}(\text{H}_2\text{O})$  MOF [46].

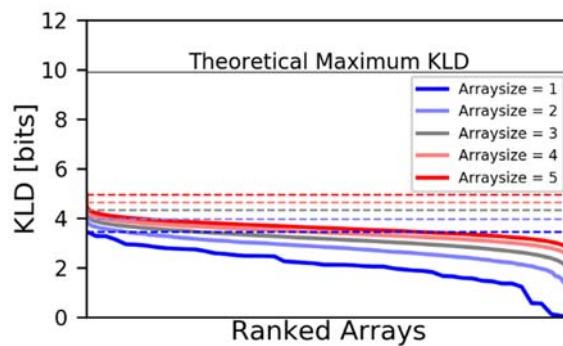


**Figure 2.** (a–c) Ternary discrete probability plots of (a) the best (Mg-MOF-74 [44]), (b) median ( $\text{Cu}_4\text{I}_4(\text{DABCO})_2$  MOF [45]), and (c) worst ( $\text{La}(\text{PODC})_{1.5}(\text{H}_2\text{O})$  MOF [46]) 1-element arrays. (d–f) Ternary discrete probability plots of the best (d) 5-, (e) 15-, and (f) 25-element arrays. Please note the change in scale of the color bar for the plots in (a–c) versus the plots in (d–f).

Given that each MOF showed selective binding for  $\text{CO}_2$  (over either  $\text{O}_2$  or  $\text{N}_2$ ), one might guess that the possibility for finding complementary combinations of MOFs to detect  $\text{O}_2$  and  $\text{N}_2$  would be impossible. Fortunately, even seemingly small differences in the adsorption behavior of the individual MOFs has a way of resolving itself into improved predictions once the arrays are large enough (see Figure 2d–f). By adding a handful of (carefully chosen) elements, we are able to narrow our prediction of all components down dramatically, and adding even more elements enables us to further improve our confidence. Figure 2d–f shows the best performing 5-, 15-, and 25- element arrays that illustrate this effect. Using only (the best-found) 5-element array, we were able to accurately predict the composition with 0.64% certainty, relative to the other 960 possible compositions; by 25-elements, we were able to accurately predict the composition of the gas with 39.3% certainty (i.e., a 39.3% chance of landing on the correct composition, 5%  $\text{CO}_2$ , 20%  $\text{O}_2$ , and 75%  $\text{N}_2$ , instead of any of the other 960 possible compositions).

### 3.2. Brute Force and Genetic Algorithm Results

As a test of the genetic algorithm, we calculated by brute force all 1-, 2-, 3-, 4-, and 5-element arrays, and then screened arrays of the same size with a genetic algorithm (see Figure 3). We ran the algorithm three times seeking the best arrays, and another three times seeking the worst arrays, for a total of 6 runs per array size. Each run of the genetic algorithm included 20 arrays per generation and 200 generations per run. As shown in Figure 3, the genetic algorithm is clearly successful in finding the highest performing arrays.

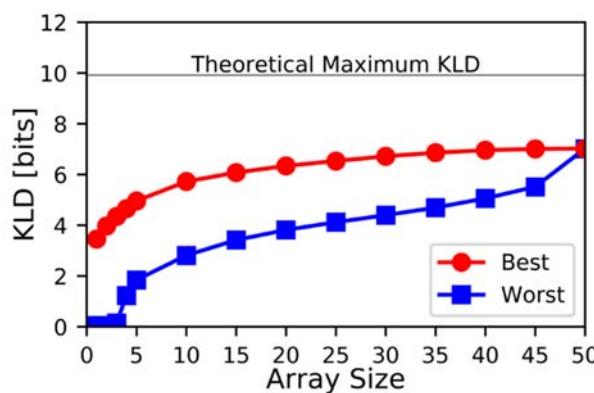


**Figure 3.** Validation of the genetic algorithm (dashed lines) against brute force screening (solid lines) for 1- to 5-element arrays. The dashed lines are the best array of each size predicted by the genetic algorithm, and the solid lines are all arrays from the brute force screening. There are more arrays as array size increases; the results are simply stretched so that the best/median/worst arrays are vertically aligned.

### 3.3. Maximum KLD vs. Number of Elements in an Array

Another key advantage of the KLD metric is that its range of values is limited only by the number of possible outcomes, so in the case of gas sensing, the number of possible compositions. As such, it can be used to compare arrays independent of size, individual sensing elements, or even the type of sensing mechanism, so long as the number of possible compositions is the same. Consequently, we can use the KLD value to show the advantage of adding more elements to a sensing array, or conversely, the diminishing returns of adding too many elements.

Figure 4 shows the best and worst KLD values of the arrays by size. Only arrays of size 1–5 were done using the brute force method; all other array sizes were screened via the genetic algorithm, hence the best, worst, and average KLD is that of all the screened arrays, not of all possible arrays.



**Figure 4.** KLD in bits of best-found and worst-found as a function of the number of elements in array.

It is clear that for smaller arrays, the overall performance is worse. The maximum achieved KLD for a 1-element sensor is only 3.46 bits, while for a two-element sensor, it is 3.97 bits, and for a three-element sensor, it is 4.35 bits. By 45 elements, the maximum-achieved KLD is 7.01 bits. It also becomes apparent from Figure 4 that the range of possible KLD values is much greater for small arrays than for large ones. Although the decrease in variability of the KLD for large arrays moderates the need for advanced screening approaches like the genetic algorithm, it underscores the advantages of employing many-element sensors over 1-element sensors.

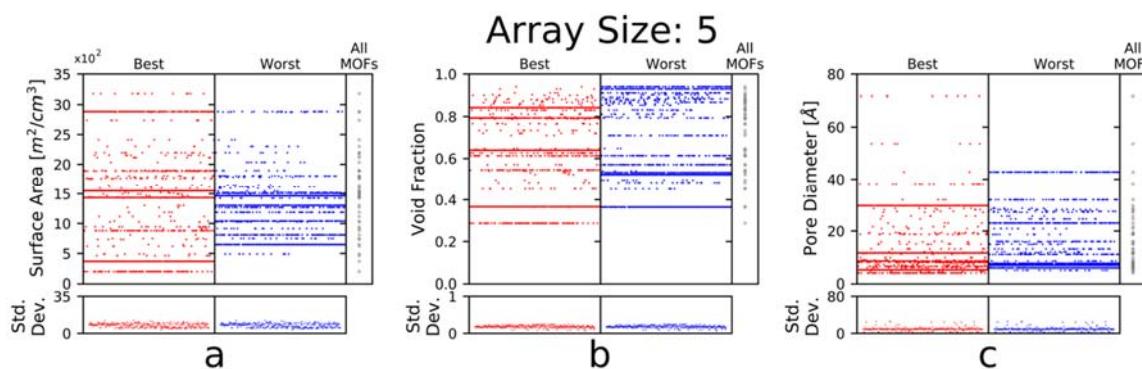
## 4. Discussion

Identifying combinations of physical properties that lead to high performance sensor arrays could give insights on how to build sensing arrays that use MOFs outside of the 50 specifically considered in

this study. Thus, here we examine the combinations of physical properties of the MOFs in the best and worst performing arrays. Intuitively, one might expect that an array made up of a diverse set of materials would perform better than an array of nearly identical materials. Conversely, it has been shown that arrays with similar materials do better at handling noise [38].

Although there are many different physical properties one could choose to examine, there are a few that are particularly likely to be important based on known trends in gas adsorption. At low pressures, adsorption behavior is typically governed by the heat of adsorption. The relative heats of adsorption between different components also provides insight into selectivity, which could be beneficial for sensing applications [35,47,48]. At moderate pressures, adsorption behavior has been shown to correlate strongly with surface area [35,47–49]. And finally, at high pressures, adsorption behavior is dominated by the free volume of the pores of the material, with gas molecules successfully packing all accessible pores and channels [35,47–49].

Hence, we examined the influence of the following properties on gas array sensing performance: volumetric surface area, void fraction, and pore diameter. In Figure 5, the 300 best and 300 worst 5-element arrays, ranked from best to worst, are plotted against the physical properties of the elements in the array. We chose to examine 5-element arrays, rather than larger arrays, since with smaller arrays it is harder to compensate for bad elements. Furthermore, with larger arrays, we observed a narrowing of the minimum and maximum KLD values. For 5-element arrays, the KLD of the 300 best-found arrays ranges from 4.94–4.58 bits, and for the 300 worst-found arrays the KLD ranges from 2.51–1.87 bits.



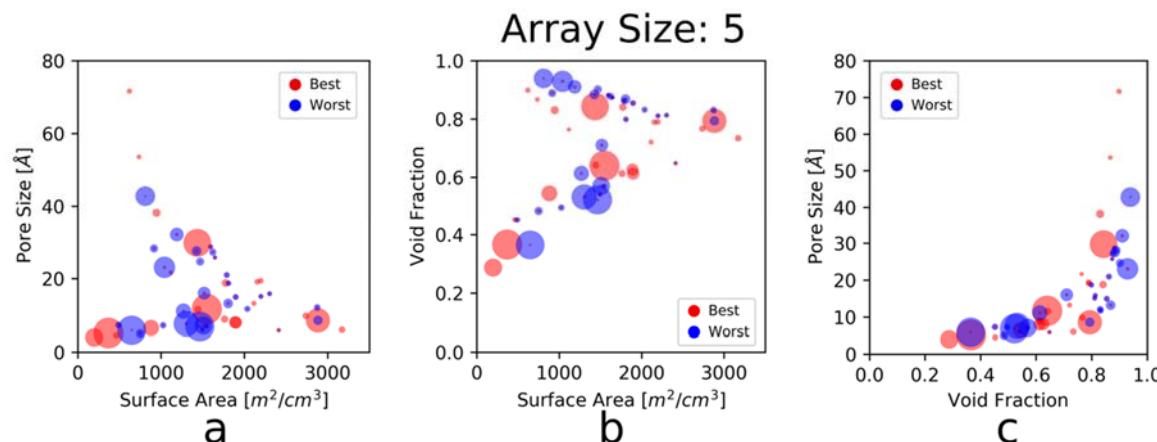
**Figure 5.** Physical property vs. rank (top/bottom 300 arrays), where (a) examines the surface area, (b) examines the void fraction, and (c) examines the pore diameter. The bottom portion of each plot shows the standard deviation of the corresponding physical property for each array (e.g., one array with the following five void fractions [0.79, 0.84, 0.54, 0.37, 0.64] would have a standard deviation in its void fraction of 0.17). The appearance of horizontal ‘lines’ corresponds to MOFs that are frequently featured in the best/worst arrays and gives us insight into both the (un)desirable properties as well as the (un)desirable spread of those properties.

With volumetric surface area, the best arrays feature elements spanning a wide range of the available surface areas, corresponding to a higher standard deviation. Conversely, the worst arrays feature elements with a narrow range of surface areas, corresponding to a lower standard deviation, indicating that for this feature, a diverse set of materials is preferred.

Similarly, the best arrays feature elements with a diverse set of void fractions. Here, however, both the best and worst arrays feature similar ranges of void fractions, and consequently the standard deviations are comparable. Nevertheless, the best arrays show more diversity over that range, whereas with the worst arrays often contain sets of elements with nearly identical void fractions, once again indicating that a diverse set of materials seems to be preferred. Additionally, the void fractions of the worst arrays are shifted higher than those for the best arrays, which is likely to lead to decreased selectivity, especially with similarly sized gases.

Consistent with before, we notice that the best arrays feature a diverse distribution of pore sizes; however, this diversity occurs over a somewhat limited range [ $\sim 4$ – $30$  Å], as compared to the worst arrays [ $\sim 5$ – $45$  Å]. This likely indicates that, as before, a diverse set of elements is beneficial, but that there is a limit to the usefulness of excessively small ( $<\sim 4$  Å) or large pores ( $>\sim 30$  Å), as beyond those limits they either exclude all gases or are so large that the volumetric density of chemically selective adsorption sites is too low for appreciably selective binding.

In addition to examining the complementarity of the individual physical properties of the sensing materials, we also wanted to examine whether there was complementarity across different physical properties (see Figure 6). For example, it might be desirable to have an element with a high surface area, but a narrow pore diameter, so that it selectively adsorbs only smaller gases, but in large quantities, providing a uniquely good signal for detecting those components. No complementary effect between any two properties (e.g., large surface areas complementing small pore sizes) becomes readily apparent, so it might be safe to conclude that for this gas mixture at these conditions, it is enough to consider only the individual properties of the elements. Nevertheless, there is still a noticeable clustering of the blue dots (worst array elements) compared to the red dots (best array elements), which continues to suggest that diversity of properties is beneficial for high-performing arrays.



**Figure 6.** Physical property coupling (top/bottom 300 arrays), where (a) examines the relationship between pore size and surface area, (b) examines the relationship between void fraction and surface area, and (c) examines the relationship between pore size and void fraction. Dot color corresponds to elements of the best/worst arrays (red is for best arrays, blue for worst arrays) and dot size corresponds to the number of times the element is featured in the best/worst arrays (a larger dot corresponds to being present in more arrays). This is the same set of arrays used in Figure 5.

When considering the data shown in Figures 5 and 6, it is important to remember that there are only 50 MOFs to choose from, and the gas mixtures feature only three gases of similar sizes (kinetic diameter of  $\text{CO}_2$  is 3.30 Å, of  $\text{O}_2$  is 3.46 Å, and of  $\text{N}_2$  is 3.64 Å) [50]. As such, the selected MOFs may not be representative of the whole material class (though they were chosen to be relatively diverse) and the gases are clearly not representative of gases species at large. Thus, the conclusions from the above figures may not be applicable for all gas-sensing applications, but rather specific to detecting  $\text{CO}_2$  in air. Moreover, with  $\text{CO}_2$  being strongly polar (relative to  $\text{N}_2$  and  $\text{O}_2$ ) and more likely to exhibit strong binding relative to these gases, it is plausible that other features, such as open metal sites, have a greater influence on the adsorption behavior, and thus are better predictors of and design criteria for the best/worst arrays than the features examined here [51–53]. Nevertheless, our study and the methodology described herein can help guide our intuition with regards to the importance of these features for gas sensing.

Given these results, we would recommend to an experimentalist designing gas-sensing arrays for  $\text{CO}_2$  in air to choose MOFs with a wide variety of surface areas and void fractions. Furthermore, we

would recommend avoiding MOFs whose pore sizes are much larger than size of the gases of interest, in order to achieve improved selectivity.

Given these insights, experimentalists aiming to develop a MOF-based sensing array, specifically for CO<sub>2</sub> sensing, should aim for a set of MOFs with diversity of features. Namely, they should choose MOFs with a wide range of surface areas. A wide range of both void fractions and pore sizes is also useful, but one should be careful to choose MOFs over a limited scope (~0.3–0.85 for void fraction and less than ~30 Å for pore size) as MOFs with too little or too much empty space will become less selective by either adsorbing no gas or all gases, respectively.

## 5. Conclusions

Herein, we have described a methodology for screening gas sensing arrays of MOFs, specifically for the detection of CO<sub>2</sub> in air. Additionally, we examined the physical properties of the MOF elements in the array to improve intuition for the design of sensing arrays.

An interesting feature of the gas mixture studied was the high sensitivity of all MOFs towards CO<sub>2</sub>. Although this result gave the impression that reliably detecting the O<sub>2</sub> and N<sub>2</sub> would be difficult, we found that with even relatively small arrays, we were able to accurately resolve their composition. Furthermore, our analysis of the physical properties of the MOFs seems to confirm our intuition that diverse arrays lead to improved sensing. A noteworthy caveat to this is that there seems to be a practical range for certain features, namely pore size, as with excessively small or large pores, selectivity is diminished by the exclusion or inclusion of all the gases present in the mixture, respectively. Although this limit is likely to change depending on the gases of interest, it seems reasonable that this concept should hold for a broad range of sensing applications.

In the future, in addition to ranking arrays over a range of compositions, we are interested in expanding the design and screening approach to include thousands of MOFs, as well as studying more complex gas mixtures, including common interferences such as carbon monoxide and water. However, this will require implementing new efficiencies to the overall procedure in order to keep computational times and costs within reason.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/1424-8220/20/3/924/s1>, Table S1: Physical Properties of MOF Structures, Table S2: Parameters of Framework Atoms, Table S3: Parameters of Gas Molecule Bodies, Table S4: Critical Parameters of Gas Molecules, Figure S1: Overview of Analysis Procedure, Figure S2: Mapping of Mass Uptake to Probability, Figure S3: Component-wise Probability, Figure S4: Example Decision Procedure for Mutation Strategy, and Figure S5: Flowchart Overview of the Genetic Algorithm.

**Author Contributions:** B.A.D. and C.E.W. conceived and designed the experiments; B.A.D. performed the experiments; B.A.D. and C.E.W. analyzed the results; B.A.D. and C.E.W. wrote the paper. All authors have read and agreed to the published version of the manuscript.

**Funding:** We are grateful for support from the National Science Foundation for funding via their EAGER program: CBET- 1937179. We are additionally grateful to the Clinical and Translational Science Institute at the University of Pittsburgh for their support via the PInCh award, and also to the University of Pittsburgh Center for Medical Innovation for providing an early stage seed grant.

**Acknowledgments:** B.A.D. and C.E.W. gratefully acknowledge support from the University of Pittsburgh's Swanson School of Engineering and the Center for Research Computing for computational resources. B.A.D. also thanks Jenna Gustafson and Paul Boone for their support regarding this work.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Nagle, H.T.; Gutierrez-Osuna, R.; Schiffman, S.S. The how and why of electronic noses. *IEEE Spectr.* **1998**, *35*, 22–31. [[CrossRef](#)]
2. Arshak, K.; Moore, E.; Lyons, G.M.; Harris, J.; Clifford, S. A review of gas sensors employed in electronic nose applications. *Sens. Rev.* **2004**, *24*, 181–198. [[CrossRef](#)]

3. Röck, F.; Barsan, N.; Weimar, U. Electronic nose: Current status and future trends. *Chem. Rev.* **2008**, *108*, 705–725. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Adiguzel, Y.; Kulah, H. Breath sensors for lung cancer diagnosis. *Biosens. Bioelectron.* **2015**, *65*, 121–138. [\[CrossRef\]](#)
5. Gardner, J.W.; Shin, H.W.; Hines, E.L. An electronic nose system to diagnose illness. *Sens. Actuators B Chem.* **2000**, *70*, 19–24. [\[CrossRef\]](#)
6. Cellini, A.; Blasioli, S.; Biondi, E.; Bertaccini, A.; Braschi, I.; Spinelli, F. Potential applications and limitations of electronic nose devices for plant disease diagnosis. *Sensors* **2017**, *17*, 2596. [\[CrossRef\]](#)
7. Kou, L.; Zhang, D.; Liu, D. A novel medical e-nose signal analysis system. *Sensors* **2017**, *17*, 402. [\[CrossRef\]](#)
8. Permentier, K.; Vercammen, S.; Soetaert, S.; Schellemans, C. Carbon dioxide poisoning: A literature review of an often forgotten cause of intoxication in the emergency department. *Int. J. Emerg. Med.* **2017**, *10*. [\[CrossRef\]](#)
9. Cavallari, M.R.; Izquierdo, J.E.E.; Braga, G.S.; Dirani, E.A.T.; Pereira-da-Silva, M.A.; Rodríguez, E.F.G.; Fonseca, F.J. Enhanced Sensitivity of Gas Sensor Based on Poly(3-hexylthiophene) Thin-Film Transistors for Disease Diagnosis and Environment Monitoring. *Sensors* **2015**, *15*, 9592–9609. [\[CrossRef\]](#)
10. Burgués, J.; Hernández, V.; Lilienthal, A.J.; Marco, S. Smelling Nano Aerial Vehicle for Gas Source Localization and Mapping. *Sensors* **2019**, *19*, 478. [\[CrossRef\]](#)
11. González-Jiménez, J.; Monroy, J.; Blanco, J.L. The Multi-Chamber Electronic Nose—An Improved Olfaction Sensor for Mobile Robotics. *Sensors* **2011**, *11*, 6145–6164. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Loutfi, A.; Coradeschi, S.; Karlsson, L.; Broxvall, M. Putting olfaction into action: Using an electronic nose on a multi-sensing mobile robot. In Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566), Sendai, Japan, 28 September–2 October 2004; Volume 1, pp. 337–342. [\[CrossRef\]](#)
13. Yamazoe, N.; Miura, N. Environmental gas sensing. *Sens. Actuators B Chem.* **1994**, *20*, 95–102. [\[CrossRef\]](#)
14. Zosel, J.; Oelßner, W.; Decker, M.; Gerlach, G.; Guth, U. The measurement of dissolved and gaseous carbon dioxide concentration. *Meas. Sci. Technol.* **2011**, *22*, 072001. [\[CrossRef\]](#)
15. Duren, R.M.; Miller, C.E. Measuring the carbon emissions of megacities. *Nat. Clim. Change* **2012**, *2*, 560–562. [\[CrossRef\]](#)
16. Oldenburg, C.M.; Unger, A.J.A. On Leakage and Seepage from Geologic Carbon Sequestration Sites. *Vadose Zone J.* **2003**, *2*, 287–296. [\[CrossRef\]](#)
17. Lewicki, J.L.; Oldenburg, C.M.; Dobeck, L.; Spangler, L. Surface CO<sub>2</sub> leakage during two shallow subsurface CO<sub>2</sub> releases. *Geophys. Res. Lett.* **2007**, *34*. [\[CrossRef\]](#)
18. Romanak, K.D.; Bennett, P.C.; Yang, C.; Hovorka, S.D. Process-based approach to CO<sub>2</sub> leakage detection by vadose zone gas monitoring at geologic CO<sub>2</sub> storage sites. *Geophys. Res. Lett.* **2012**, *39*. [\[CrossRef\]](#)
19. Srisont, S.; Chirachariyavej, T.; Peonim, A.V.M.V. A carbon dioxide fatality from dry ice. *J. Forensic Sci.* **2009**, *54*, 961–962. [\[CrossRef\]](#)
20. Satish, U.; Mendell, M.J.; Shekhar, K.; Hotchi, T.; Sullivan, D.; Streufert, S.; Fisk, W.J. Is CO<sub>2</sub> an indoor pollutant? Direct effects of low-to-moderate CO<sub>2</sub> concentrations on human decision-making performance. *Environ. Health Perspect.* **2012**, *120*, 1671–1677. [\[CrossRef\]](#)
21. Pandey, S.K.; Kim, K.-H. The Relative Performance of NDIR-based Sensors in the Near Real-time Analysis of CO<sub>2</sub> in Air. *Sensors* **2007**, *7*, 1683–1696. [\[CrossRef\]](#) [\[PubMed\]](#)
22. MedlinePlus—Health Information from the National Library of Medicine. Available online: <https://medlineplus.gov/> (accessed on 8 October 2019).
23. Campbell, M.G.; Dincă, M. Metal–organic frameworks as active materials in electronic sensor devices. *Sensors* **2017**, *17*, 1108. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Gustafson, J.A.; Wilmer, C.E. Intelligent selection of metal–organic framework arrays for methane sensing via genetic algorithms. *ACS Sens.* **2019**, *4*, 1586–1593. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Gustafson, J.A.; Wilmer, C.E. Computational design of metal–organic framework arrays for gas sensing: Influence of array size and composition on sensor performance. *J. Phys. Chem. C* **2017**, *121*, 6033–6038. [\[CrossRef\]](#)
26. Gustafson, J.A.; Wilmer, C.E. Optimizing information content in MOF sensor arrays for analyzing methane-air mixtures. *Sens. Actuators B Chem.* **2018**, *267*, 483–493. [\[CrossRef\]](#)

27. Sturluson, A.; Sousa, R.; Zhang, Y.; Huynh, M.T.; Laird, C.; York, A.H.P.; Silsby, C.; Chang, C.-H.; Simon, C. Curating metal-organic frameworks to compose robust gas sensor arrays in dilute conditions. *ACS Appl. Mater. Interfaces* **2020**, *12*, 6546–6564. [\[CrossRef\]](#)

28. Burrows, A.D.; Frost, C.G.; Mahon, M.F.; Richardson, C. Post-synthetic modification of tagged metal-organic frameworks. *Angew. Chem. Int. Ed.* **2008**, *47*, 8482–8486. [\[CrossRef\]](#)

29. Farha, O.K.; Eryazici, I.; Jeong, N.C.; Hauser, B.G.; Wilmer, C.E.; Sarjeant, A.A.; Snurr, R.Q.; Nguyen, S.T.; Yazaydin, A.Ö.; Hupp, J.T. Metal-organic framework materials with ultrahigh surface areas: Is the Sky the limit? *J. Am. Chem. Soc.* **2012**, *134*, 15016–15021. [\[CrossRef\]](#)

30. Kreno, L.E.; Leong, K.; Farha, O.K.; Allendorf, M.; Van Duyne, R.P.; Hupp, J.T. Metal-organic framework materials as chemical sensors. *Chem. Rev.* **2012**, *112*, 1105–1125. [\[CrossRef\]](#)

31. Zhu, X.; Zheng, H.; Wei, X.; Lin, Z.; Guo, L.; Qiu, B.; Chen, G. Metal-organic framework (MOF): A novel sensing platform for biomolecules. *Chem. Commun.* **2013**, *49*, 1276. [\[CrossRef\]](#)

32. Zhou, J.; Kitagawa, S. Metal-organic frameworks (MOFs). *Chem. Soc. Rev.* **2014**, *43*, 5415–5418. [\[CrossRef\]](#) [\[PubMed\]](#)

33. Zhai, Q.-G.; Bu, X.; Mao, C.; Zhao, X.; Daemen, L.; Cheng, Y.; Ramirez-Cuesta, A.J.; Feng, P. An ultra-tunable platform for molecular engineering of high-performance crystalline porous materials. *Nat. Commun.* **2016**, *7*, 13645. [\[CrossRef\]](#) [\[PubMed\]](#)

34. Maurin, G.; Serre, C.; Cooper, A.; Férey, G. The new age of MOFs and of their porous-related solids. *Chem. Soc. Rev.* **2017**, *46*, 3104–3107. [\[CrossRef\]](#) [\[PubMed\]](#)

35. Düren, T.; Bae, Y.-S.; Snurr, R.Q. Using molecular simulation to characterise metal-organic frameworks for adsorption applications. *Chem. Soc. Rev.* **2009**, *38*, 1237–1247. [\[CrossRef\]](#)

36. Getman, R.B.; Bae, Y.-S.; Wilmer, C.E.; Snurr, R.Q. Review and analysis of molecular simulations of methane, hydrogen, and acetylene storage in metal-organic frameworks. *Chem. Rev.* **2012**, *112*, 703–723. [\[CrossRef\]](#)

37. Zeitler, T.R.; Allendorf, M.D.; Greathouse, J.A. Grand canonical Monte Carlo simulation of low-pressure methane adsorption in nanoporous framework materials for sensing applications. *J. Phys. Chem. C* **2012**, *116*, 3492–3502. [\[CrossRef\]](#)

38. Fernández Romero, L. Understanding the Role of Sensor Diversity and Redundancy to Encode for Chemical Information in Gas Sensor Arrays. Ph.D. Thesis, Universitat de Barcelona, 2016.

39. Dubbeldam, D.; Calero, S.; Ellis, D.E.; Snurr, R.Q. RASPA: Molecular simulation software for adsorption and diffusion in flexible nanoporous materials. *Mol. Simul.* **2016**, *42*, 81–101. [\[CrossRef\]](#)

40. Chung, Y.G.; Camp, J.; Haranczyk, M.; Sikora, B.J.; Bury, W.; Krungleviciute, V.; Yildirim, T.; Farha, O.K.; Sholl, D.S.; Snurr, R.Q. Computation-ready, experimental metal-organic frameworks: A tool to enable high-throughput screening of nanoporous crystals. *Chem. Mater.* **2014**, *26*, 6185–6192. [\[CrossRef\]](#)

41. Wilmer, C.E.; Kim, K.C.; Snurr, R.Q. An extended charge equilibration method. *J. Phys. Chem. Lett.* **2012**, *3*, 2506–2511. [\[CrossRef\]](#)

42. Martin, M.G.; Siepmann, J.I. Transferable potentials for phase equilibria. 1. United-atom description of n-alkanes. *J. Phys. Chem. B* **1998**, *102*, 2569–2577. [\[CrossRef\]](#)

43. MacKay, D.J.C.; Kay, D.J.C.M. *Information Theory, Inference and Learning Algorithms*; Cambridge University Press: Cambridge, UK, 2003; ISBN 978-0-521-64298-9.

44. Rosi, N.L.; Kim, J.; Eddaoudi, M.; Chen, B.; O’Keeffe, M.; Yaghi, O.M. Rod packings and metal-organic frameworks constructed from rod-shaped secondary building units. *J. Am. Chem. Soc.* **2005**, *127*, 1504–1518. [\[CrossRef\]](#) [\[PubMed\]](#)

45. Braga, D.; Maini, L.; Mazzeo, P.P.; Ventura, B. Reversible interconversion between luminescent isomeric metal-organic frameworks of  $[Cu_4I_4(DABCO)_2]$  (DABCO=1,4-Diazabicyclo [2.2.2]octane). *Chem. Eur. J.* **2010**, *16*, 1553–1559. [\[CrossRef\]](#) [\[PubMed\]](#)

46. Zhuang, G.-L.; Kong, X.-J.; Long, L.-S.; Huang, R.-B.; Zheng, L.-S. Effect of lanthanide contraction on crystal structures of lanthanide coordination polymers with 2,5-piperazinedione-1,4-diacetic acid. *CrystEngComm* **2010**, *12*, 2691–2694. [\[CrossRef\]](#)

47. Frost, H.; Düren, T.; Snurr, R.Q. Effects of surface area, free volume, and heat of adsorption on hydrogen uptake in metal-organic frameworks. *J. Phys. Chem. B* **2006**, *110*, 9565–9570. [\[CrossRef\]](#)

48. Yang, Q.; Zhong, C. Molecular Simulation of Carbon Dioxide/Methane/Hydrogen Mixture Adsorption in Metal-Organic Frameworks. *J. Phys. Chem. B* **2006**, *110*, 17776–17783. [\[CrossRef\]](#)

49. Li, J.-R.; Ma, Y.; McCarthy, M.C.; Sculley, J.; Yu, J.; Jeong, H.-K.; Balbuena, P.B.; Zhou, H.-C. Carbon dioxide capture-related gas adsorption and separation in metal-organic frameworks. *Coord. Chem. Rev.* **2011**, *255*, 1791–1823. [[CrossRef](#)]
50. Ismail, A.F.; Khulbe, K.C.; Matsuura, T. Fundamentals of gas permeation through membranes. In *Gas Separation Membranes: Polymeric and Inorganic*; Ismail, A.F., Chandra Khulbe, K., Matsuura, T., Eds.; Springer International Publishing: Cham, Switzerland, 2015; pp. 11–35. ISBN 978-3-319-01095-3. [[CrossRef](#)]
51. Poloni, R.; Lee, K.; Berger, R.F.; Smit, B.; Neaton, J.B. Understanding trends in CO<sub>2</sub> adsorption in metal-organic frameworks with open-metal sites. *J. Phys. Chem. Lett.* **2014**, *5*, 861–865. [[CrossRef](#)]
52. Gardner, J.W.; Boilot, P.; Hines, E.L. Enhancing electronic nose performance by sensor selection using a new integer-based genetic algorithm approach. *Sens. Actuators B Chem.* **2005**, *106*, 114–121. [[CrossRef](#)]
53. Perry, J.J.; Teich-McGoldrick, S.L.; Meek, S.T.; Greathouse, J.A.; Haranczyk, M.; Allendorf, M.D. Noble Gas Adsorption in Metal–Organic Frameworks Containing Open Metal Sites. *J. Phys. Chem. C* **2014**, *118*, 11685–11698. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).