A Temporal Filter to Extract Doped Conducting Polymer Information Features from an Electronic Nose

Wiem Haj Ammar, Aicha Boujnah, Antoine Baron, Aimen Boubaker, Adel Kalboussi, Kamal Lmimouni, Sébastien Pecqueur

electronics

Article

A Temporal Filter to Extract Doped Conducting Polymer Information Features from an Electronic Nose

Wiem Haj Ammar [1], Aicha Boujnah [1], Antoine Baron [2], Aimen Boubaker [1], Adel Kalboussi [1], Kamal Lmimouni [2] and Sébastien Pecqueur [2,*]

[1] Department of Physics, University of Monastir, Monastir 5000, Tunisia; wiem.hajammar@fsm.rnu.tn (W.H.A.); aicha.boujnah@fsm.rnu.tn (A.B.); aimen.boubaker@ipein.rnu.tn (A.B.); adel.kalboussi@fsm.rnu.tn (A.K.)
[2] University of Lille, CNRS, Centrale Lille, University of Polytechnique Hauts-de-France, UMR 8520-IEMN, F-59000 Lille, France; antoine.baron@univ-lille.fr (A.B.); kamal.lmimouni@univ-lille.fr (K.L.)
* Correspondence: sebastien.pecqueur@iemn.fr

Abstract: Identifying relevant machine learning features for multi-sensing platforms is both an applicative limitation to recognize environments and a necessity to interpret the physical relevance of transducers' complementarity in their information processing. Particularly for long acquisitions, feature extraction must be fully automatized without human intervention and resilient to perturbations without significantly increasing the computational cost of a classifier. In this study, we investigate the relative resistance and current modulation of a 24-dimensional conductimetric electronic nose, which uses the exponential moving average as a floating reference in a low-cost information descriptor for environment recognition. In particular, we identified that depending on the structure of a linear classifier, the 'modema' descriptor is optimized for different material sensing elements' contributions to classify information patterns. The low-pass filtering optimization leads to opposite behaviors between unsupervised and supervised learning: the latter favors longer integration of the reference, allowing the recognition of five different classes over 90%, while the first one prefers using the latest events as its reference to cluster patterns by environment nature. Its electronic implementation shall greatly diminish the computational requirements of conductimetric electronic noses for on-board environment recognition without human supervision.

Keywords: conducting polymers; electronic noses; exponential moving average; sensing; classifiers

Citation: Haj Ammar, W.; Boujnah, A.; Baron, A.; Boubaker, A.; Kalboussi, A.; Lmimouni, K.; Pecqueur, S. A Temporal Filter to Extract Doped Conducting Polymer Information Features from an Electronic Nose. *Electronics* **2024**, *13*, 497. https://doi.org/10.3390/electronics13030497

Academic Editor: Chiman Kwan

Received: 22 December 2023
Revised: 19 January 2024
Accepted: 22 January 2024
Published: 24 January 2024

1. Introduction

Consumer electronics have experienced a revolution in the use of sensing hardware as machine learning-supported information generators, thanks to the Internet of Things [1]. While sensors were used exclusively for metrology, sensing devices are now exploited to increase the perception field of electronics with environmentally sensitive input layers, constantly generating new sets of organic data to be classified. If smartphones and watches have now become the epidermal nerves of our modern society, it is at the cost of valuing hardware more for correlating generated information than quantifying physics as metrological devices [2,3]. Microphones, photodetectors and accelerometers are being used less to measure harmonics, photons or displacements than to recognize spoken queries, heartbeats or gesture patterns from large populations at a given time, so the underlying physical mechanisms ruling their functioning may appear quite irrelevant for an application. However, there is still a large part of the information we sense with our biology that cannot be classified artificially with hardware: this is the case with olfacto-mimetics [4]. Smells are less something to be measured than to be classified [5]. Fragrances of volatile molecules trigger reversible changes of states on our G-coupled olfacto-receptor proteins, for which the imprint encodes a specific class of stimuli in our brain, among a trillion that could be recognized [6]. These vast combinations of smells can hardly be mapped in a low-dimensional space [7–9]. Moreover, the complexity of the interaction between

one odoriferous molecule with a single human receptor has only recently been experimentally characterized [10]. Pursuing such tremendous work at the level of all possible volatile molecules with all possible protein receptors to decipher such a perception space as olfaction seems unrealistic. However, considering individual molecules as "primary scents" (as we do with colors) seems quite irrelevant to constituting groups of patterns to break down the chemical compositions of smells. An odoriferous pattern has multiple identities as a chemical composition, as similar patterns can be induced by very different molecules [11–13], and molecules that are structurally very close to one another can induce very different perception patterns [12,14–18]. Identifying the right sets of material sensitizers that embed enough orthogonality in their information features is key to co-integrating electronic noses (eNoses) that perceive enough differences in volatile molecule imprinting [19–21]. Conducting polymer materials have already been extensively studied as sensitive electro-transducers for electronic noses because of their versatility, behaving as semiconductors as well as chemically specific macromolecular probes [22–24]. As semiconductors, conducting polymers need to be doped by adding electron acceptors or donors to efficiently conduct electricity [25–27]. Former studies have shown that varying the p-dopant of the same conducting polymer allows the tuning of the electro-transduction of different vapors on polymer semiconductors exposed in a blow [28]. Various combinations of doped conducting polymers can induce very different sensitivities to an array. And as the size of an array conditions the complexity of the classification, a careful selection of sensitive materials shall be made. The recognition rate of the array is a strong criterion that determines the material selection, which can be conditioned by various parameters. First, the recognition rate of an eNose composing these doped polymers highly depends on the software architecture of a classifier: the same dataset can either be optimally recognized in a supervised learning scheme such as LDA (linear discriminant analysis) or far less in the case of unsupervised classification by PCA (principal component analysis) or kNN (k-nearest neighbor) [29]. Second, the performance of the same classifier fed with the same raw dataset also greatly depends on the information feature that is input [30]. Different features that are representative of either the dedoping kinetics or their thermodynamics can promote recognition with different subsets of dynamical data, depending on which physical mechanisms the selected information feature promotes best [30]. Thus, preprocessing data before classification is of great importance, in order for all complementarities of a material set to constructively express themselves within a generic information feature. The association of (1) the right set of sensitive materials, (2) a software classifier architecture and (3) appropriate conditioning of the raw feature also greatly relies on (4) how systematic data can be sampled to train a system to recognize high-dimensional patterns, both for supervised schemes to train the system with enough sample and for unsupervised learning for the system to have a reliable picture of the different class organization in a high-dimensional space. To do so, the information feature to be input into the system must be simple enough to lower the computational cost for data preprocessing prior to classification. Real-time inference of an eNose is of prior importance in different applications where they show their true relevance, such as for instant detection of environmental pollutants and chemical hazards [31–33], high-throughput food-quality assessment [34–36], and their compact integration into smartphones for disease early diagnosis [37–39]. In most of these applications, a rapid response of the system is often required without users being able to routinely declare when samples of interest are being exposed. In cases where information features require some parametrizing by an external operator, it limits the approach to small set of data and requires experimentalists to manually process training sets and parametrize the features. This is the case for most features used for conductimetric eNoses, such as the resistance variation or its integration over time before or after a sample exposure at t_0 [40–42]. These features require discretizing time sequences to elementary samples by defining t_0 for each sample as setup-dependent external parameters. For the relative resistance modulation, declaring t_0 is required to define a floating reference resistance R_0 as $R(t_0)$ for each sensing unit measured in a reference environment. This reference has to be

periodically renewed and routinely defined when opening to a class-specific environment and when a sensing element reaches its steady state, measuring $R(t_0 + t)$. Both parameters are required to routinely evaluate the information feature $(R(t_0 + t) - R_0)/R_0$ and are highly sensitive to the R_0 values, which are artificially set by an experimental user upon exposures to reference environments (purges) and class-specific environments. Also, such features are inadequate when periodic referencing before each class exposure cannot be implemented experimentally: for instance, when training an electronic nose on learning pollution peaks in an outdoor test environment. For applications where discretizing the time series by samples is doable (such as for assessing the quality of items), delays between the actual exposures and the setting of t_0 can bias the feature and alter the classification.

The physical triviality of an information pattern does not condition the recognition performances of a classifier algorithm; therefore, features do not necessarily have to be interpreted ab initio from physical mechanisms prior to being used as information descriptors in machine learning. Therefore, one may mathematically build indicators that may not be a priori representative of a particular physical property. This is the case in studies led by Vergara, Huerta and coworkers [43,44], where the exponential moving average (ema) of resistance is used as a floating reference in the information feature of sensing elements' responses in a conductimetric electronic nose classifier. In their work, the choice for such an information feature was motivated by its extensive use in the predictive analysis of dynamical signals where the low explicability of governing physical mechanisms does not hinder classification, such as in financial analysis. Thanks to this strategy, the intervention of a user to set t_0 is not required. At any time during the measurement of a dynamical vector $X(t)$, a floating reference $ema_\alpha(X(t))$ can easily be defined as Equations (1) and (2):

$$ema_\alpha(X(t)) = \alpha \cdot X(t) + (1 - \alpha) \cdot ema_\alpha(X(t - \Delta T)) \quad (1)$$

$$ema_\alpha(X(t = 0)) = X(t = 0) \quad (2)$$

where $0 \leq \alpha \leq 1$ is the signal attenuation coefficient (or smoothing factor), here defined as a time constant, and ΔT is the sampling period of X, here set to one second. By its recursive definition, $ema_\alpha(X(t))$ is a low-cost floating reference that is defined by attenuating a fraction of the signal $X(t)$ with the former floating reference $ema_\alpha(X(t - \Delta t))$, which can be erased from memory once the feature is computed.

We present in this study its implementation in a conductimetric electronic nose using doped conducting polymers as chemo-sensitive electro-transducer materials to evaluate the $modema_\alpha(X(t))$ features as information descriptors, such as in Equation (3):

$$modema_\alpha(X(t)) = \frac{X(t) - ema_\alpha(X(t))}{ema_\alpha(X(t))} \quad (3)$$

Also, since the ema transform is a digital low-pass filter, its physical implementation in an RC-based neuron model as a circuit is discussed to optimize the design of electronic nose information generators, allowing it to behave more closely to the way an actual olfactory receptor neuron preprocesses information prior to its classification.

2. Experimental Section

Device Fabrication: The sensing hardware microfabrication is fully described in a former study [28]. Concentric Au microelectrodes (28 µm in diameter, channel length L = 400 nm, spiral electrode length W such that $W/L > 10^3$) are lithographically patterned by an e-beam in a cleanroom environment. Different materials are subsequently added to the different structures as clusters of 16 individual elements. The different 10 mg/mL formulations are deposited individually on the different clusters of electrodes by drop casting, from a solution of P3HT in its pristine state (mildly doped by O_2 as stored in air) and metal trifluorosulfonates (salts of Fe^{III}, Bi^{III}, Cu^{II}, Al^{III}, In^{III}, Dy^{III} and Ce^{III}) acting as dopants (see Figure 1).

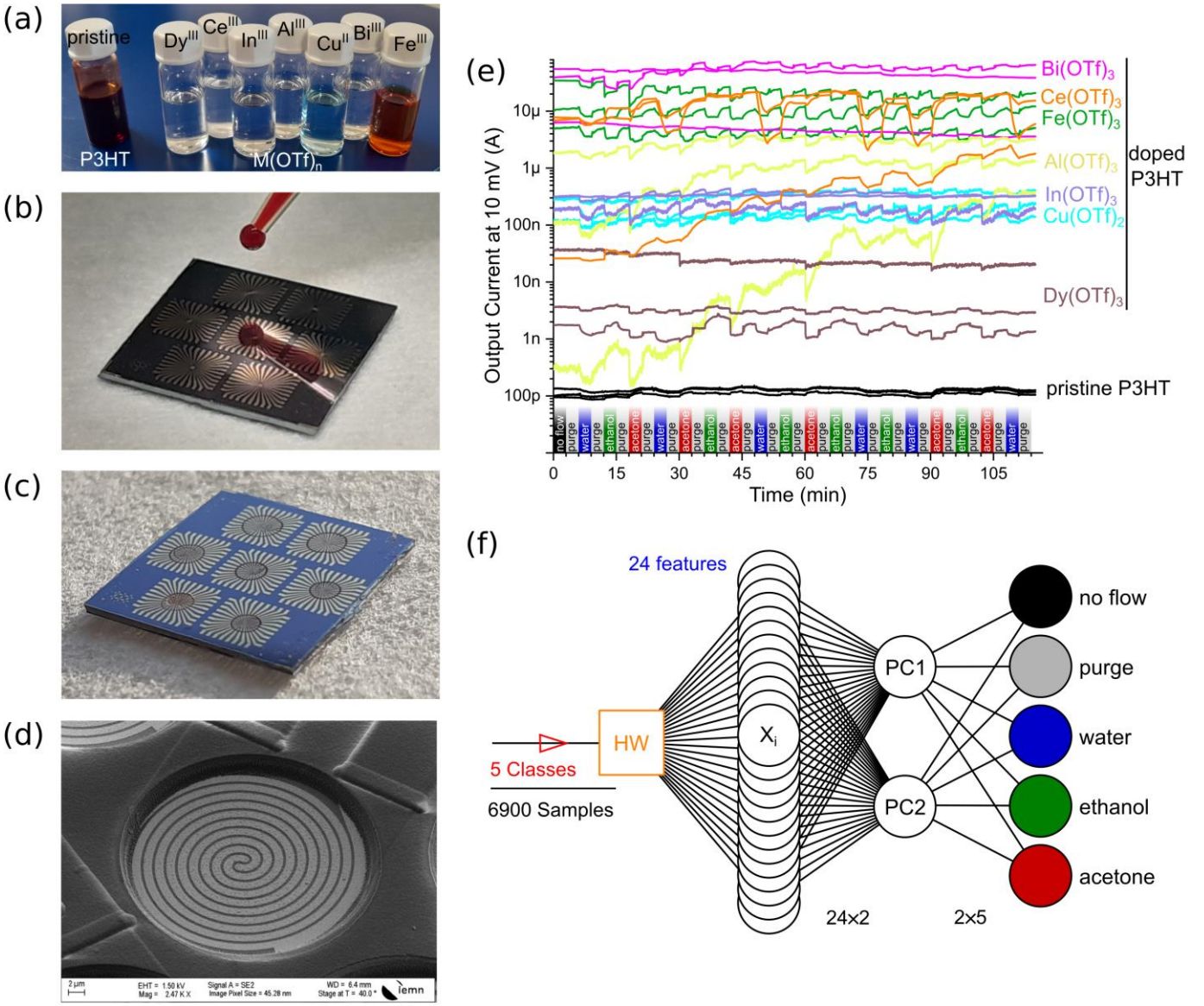

Figure 1. Unsupervised classifier for solvent vapor recognition. (**a**–**d**) Hardware used in the classification. (**a**) Photographs of the solutions used to sensitize the different clusters of conductimetric elements. (**b**) Deposition of the chemosensitive materials on the individual clusters by two sequential drop castings. (**c**) Seven clusters sensitized with different metal triflate-doped and undoped P3HT coatings as chemospecific transducers of environmental information. (**d**) Scanning electron microscope picture of the micro-electrodes of an individual sensing element used in this study. (**e**) Raw current signals extracted from a population of 24 different sensing elements supplied under DC voltage bias, coated with eight different metal triflate-doped and undoped P3HT coatings (3 sensing elements per cluster) and exposed under repeated volatile solvent vapor exposures. (**f**) Software classifier structure used in this study for unsupervised classification of the five different classes of exposure, using the first two principal components of PCA. Classification of samples measured upon exposure to water (in blue), acetone (in red) or ethanol (in green); upon purging the system (in grey); or upon starting the measurements without air flow (in black).

Electrical Characterization: The currents were measured versus time with an Agilent 4155 parameter analyzer (from Agilent Technologies, USA). Except for the first three minutes of acquisition, the sensing elements are measured in an air blow that passes through a solvent-containing vial. As the materials show great sensitivity to the analyte flow, samples were steadily exposed to 1 mL/s without interruption between samples

by permuting analysts with a bypass (more details are provided in former studies) [28]. The setup was considered under standard conditions of temperature and pressure in an indoor laboratory environment (no metrological instruments were integrated to label the environments with pressure, temperature or hygrometry readout). Solutions were tested as pure and are expected to be blown on the device in a gas phase at the partial pressure of the analytes, which equals their saturation pressure at 1 atm and 20 °C (0.24, 0.06 and 0.02 atm, respectively, for acetone, ethanol and water) [45]. The control of the different exposures, each lasting three minutes, was operated manually. Acetone-, ethanol- and water-loaded flows were exposed six times successively with permutations to avoid a systematic biasing of the data due to the analyte sequence order. Periods of purges (via an unloaded vial) separated each gas exposure to avoid cross-contaminations. The total acquisition consisted of 6900 datapoints sampled every second, constituting five classes of environments: "no flow" during the initial three minutes (180 points), "acetone", "water", "ethanol" (each of these 6 × 180 points) and "purge" between each exposure (19 × 180 + 160 points). The choice of testing the different analytes was motivated by their ease of being handled as solutions of low-hazard, high-volatility and generic molecular compounds. Each of them is a simple molecule exclusively composed of carbon, oxygen and hydrogen atoms composing different chemical functions, which induce different coordination with the oxophilic metal center of Lewis acidic p-dopants (see former studies) [28]. Discriminating ketone/alcohol rates in moist environments has various applicative interests for quality monitoring during fermentation [46] or electrocatalysis [47].

Data Analysis: Raw currents are generated and used without filtering. Sensors are considered ohmic resistors, whose resistance values are estimated from the raw current trace measured under a constant bias at 10 mV. Raw current curves have already been published [28] and are available on a public repository. This study used two linear software classifiers: PCA was used by means of the Clustvis online open-access tool [48], scaled by unit variance and computed by singular value decomposition. All PCA data are available as Supplementary Materials. The Moore–Penrose pseudo-inverse supervised classifier was adapted from a former study [49] to five-class recognition as follows:

The classifier output $Y(t) \in \mathbb{R}^m$ (for $m = 5$ classes) of a given input vector $X(t) \in \mathbb{R}^n$ from the test database (for $n = 24$ sensing units) is determined from the weight matrix $\boldsymbol{W} \in \mathbb{R}^{m \times n}$ in Equation (4):

$$Y(t) = f \circ g(\boldsymbol{W} \cdot X(t)) \tag{4}$$

The weight matrix $\boldsymbol{W}$ is computed from the pseudo-inverse matrix $\boldsymbol{X}^+ \in \mathbb{R}^{n \times p}$ for a given set of p input vectors $X(t)$ from the learning database, each of them associated with an output vector $Y(t)$ defined as a vector from the standard basis $(e_i)_{i=1}^m$ of $\mathbb{R}^m$ for each of the m classes. For p vectors from the learning database, the output matrix $\boldsymbol{Y} \in \mathbb{R}^{p \times m}$ is defined for supervised learning, so $\boldsymbol{W}$ is defined as Equation (5):

$$\boldsymbol{W} = (\boldsymbol{X}^+ \cdot \boldsymbol{Y})^{\boldsymbol{T}} \tag{5}$$

The Moore–Penrose pseudo-inverse matrix $\boldsymbol{X}^+$ is calculated for a given set of p input vectors $X(t)$, for which the input matrix $\boldsymbol{X} \in \mathbb{R}^{p \times n}$ is defined, so that Equation (6) is as follows:

$$\boldsymbol{X}^+ = \left(\boldsymbol{X}^{\boldsymbol{T}} \cdot \boldsymbol{X}\right)^{-1} \cdot \boldsymbol{X}^{\boldsymbol{T}} \tag{6}$$

The activation function $(f \circ g)$ adjusts the output $Y(t)$ sensitivity to each class density in the learning database by the g function, and input vectors are chosen randomly in the learning database. The g function is defined by the diagonal matrix $\boldsymbol{D} \in \mathbb{R}^{m \times m}$ as Equations (7) and (8):

$$g(\boldsymbol{W} \cdot X(t)) = \boldsymbol{D} \cdot \boldsymbol{W} \cdot X(t) \tag{7}$$

$$\boldsymbol{D} = diag(x_i, \ldots, x_m) \tag{8}$$

where x_i is defined as the ratio of the number of vectors that do not belong to the i^{th} class in the learning database, over m. The f function applies a "winner takes all" normalization of the $g(\boldsymbol{W}\cdot X(t)) \in \mathbb{R}^m$ vector, such that the attributed class is defined by the coordinate of $g(\boldsymbol{W}\cdot X(t))$, which has the highest value in Equations (9)–(11):

$$\forall\, i \leq m,\ \exists\, i_0(t) \leq m \quad | \quad g(\boldsymbol{W}\cdot X(t))_{i=i_0(t)} = \max(g(\boldsymbol{W}\cdot X(t))_{i=1}^{m}) \tag{9}$$

$$if\ i = i_0(t),\ then \quad f \circ g(\boldsymbol{W}\cdot X(t))_i = 1,\ else \quad f \circ g(\boldsymbol{W}\cdot X(t))_i = 0 \tag{10}$$

$$\Leftrightarrow \quad Y(t) = e_{i=i_0(t)} \tag{11}$$

3. Results

3.1. Data Collection and Feature Extraction

The hardware part of the classifier used in this study is composed of several clusters of conductimetric elements, with the metrological instrument that allows the outputting of the current $i(t)$ for each sensing element polarized under a steady voltage bias in a chemically varying environment [28]. Each cluster of conductimetric elements was sensitized with a conducting material and a dopant (see Figure 1a–d). The dopants transduce to the conducting polymer, specifically the chemical interactions they have with environmentally present volatile molecules. When exposed to an air blow, dopants imprint on each polymer, coating the presence of volatile molecules in the carrier flow by the transience of the electric current that outputs under a steady voltage (see Figure 1e). The imprint for each volatile molecule on each sensing element is specific to the doping element [28].

When acetone, water and ethanol are present in a steady air blow on the sensing array, decreases or increases in the current arise specifically on the different elements. The responses of three sensing elements coated with the same materials do not exhibit electrical properties without dispersion in current magnitude at rest and/or in current modulation during exposures. These raw data dispersions are attributed to the limitation of the deposition technique resulting in an inherent spatial heterogeneity of the materials, which conditions the reproducibility of the device performances [28]. Despite the fact that some devices even displayed significant drifts over time (see Figure 1e for the $Al(OTf)_3$-doped P3HT traces for an example), the principal component analysis (PCA) performed on the current modulation shows that most of the information is rather conditioned by the deposited materials. To classify the five different classes of exposures that are presented to the sensing layer of the hardware ("no flow", "purge", "acetone", "ethanol" and "water"), the first two principal components were used to observe whether the information feature referenced by the ema floating point could be classified in an unsupervised way (see Figure 1f). By clustering most of the data by the nature of different classes that are exposed, and by the nature of the materials that compose the two first principal components, the aim is to demonstrate whether an ema-referenced information feature can be generically used as a descriptor to recognize dynamical environments with a conducting polymer electronic nose (see Figure 2).

3.2. Exponential Moving Average as a Floating Reference

As an inverse of each other, resistances and admittances are not linearly correlated. Thus, as a linear classifier, a PCA based on one will differ from a PCA based on the other. Therefore, the ema-referenced feature study was performed in parallel with the resistance $R(t)$ (in MΩ) and the current $i(t)$ (in nA) for each sensing element as raw data vectors $X(t)$ (see Figure 2). Figure 2a,d display the typical response of a sensing element without significant drift. In Figure 2b,e, the ema referencing behaves as a filter that reduces the signal noise for low $1/\alpha^-$ values. For values of $1/\alpha^-$ typically above 30, the filtering is such that the $ema_\alpha(X(t))$ presents a noticeable temporal delay compared to the signal $X(t)$, both for the resistance and the current.

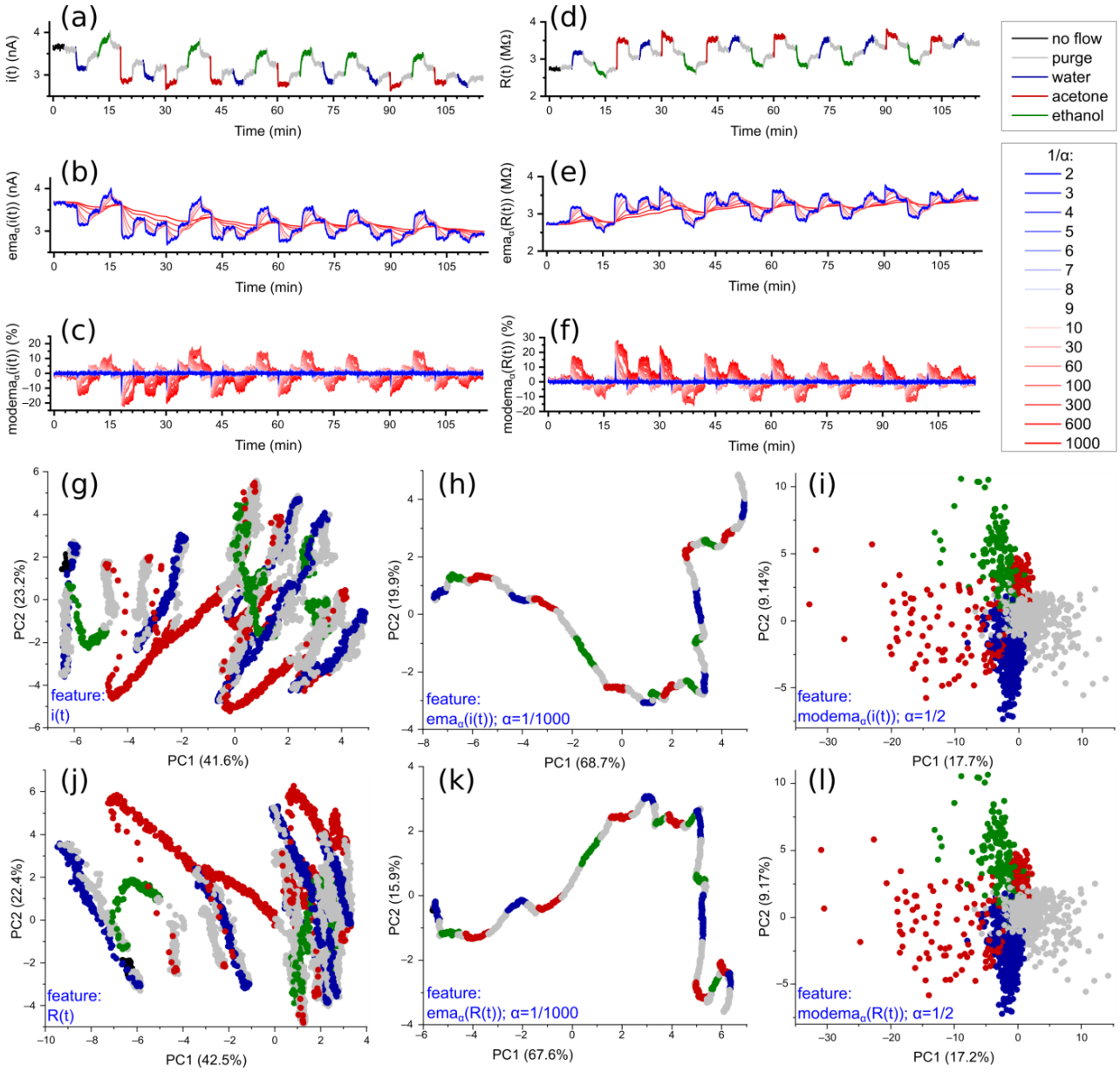

Figure 2. The *modema*() function as a transform to condition the raw data signal as relevant information descriptor for classification. (**a**,**d**) Raw data signal for one sensing element exposed under the different sequences, where each exposure is marked by a specific color in the single measurement trace of current (**a**) or resistance (**d**). (**b**,**e**) Exponential moving average (ema) transform of the raw current (**b**) and resistance (**e**) signals displayed in a and d graphs. The effect of the α parameter in the $ema_\alpha()$ function is displayed in the color gradient. (**c**,**f**) Modulation centered on the exponential moving average (modema) transform of the raw current (**b**) and resistance (**e**) signals displayed in a and d graphs. The effect of the α parameter in the $modema_\alpha()$ function is displayed in the color gradient. (**g**–**l**), First two component scores of the principal component analysis for the complete 24-dimensional datasets of different information features. Scatter colors according to the legend in (**a**,**d**). (**g**,**j**) (PC1;PC2) planes for the raw data information $i(t)$ and $R(t)$, showing poor data separation among the five environmental classes. (**h**,**k**) (PC1;PC2) planes showing data separation of the $ema_\alpha()$-transformed $i(t)$ and $R(t)$ signals (for $\alpha = 1/1000$), but poor environmental specificity. (**i**,**l**) (PC1;PC2) planes showing data separation of the $modema_\alpha()$-transformed $i(t)$ and $R(t)$ signals (for $\alpha = 1/2$) and environmental specificity for at least four different classes by nature of the exposure (as colors are no longer isolated and regrouped in four specific zones in the diagram).

This delay tends to increase with a lower value of α, where one can observe that the lower the α, the less sensitive to upcoming exposures it becomes, leaving only a slow decay representing the sensing elements' signal drift. As the smoothing quality depends highly on the signal modulation upon exposing different classes, the optimization of the filter cut-off shall be sensing element-dependent. More particularly, a sensing element that displays only moderate signal changes may result in an $ema_\alpha(X(t))$ reference that is representative of the sensing element's temporal drift, while a sensing element displaying large signal modulations may result in an $ema_\alpha(X(t))$ featuring a noise-free image of the transient signal at a given value of α. Therefore, it is expected that a sensing array showing large dispersion in signal modulations will have a much higher sensitivity when optimizing the α attenuation coefficient of an $ema_\alpha(X(t))$ floating reference than a sensing array composed of sensing elements that modulate with the same amplitudes. The sensitivity of the $ema_\alpha(X(t))$ floating reference to the α attenuation coefficient consequently affects the sensitivity of the $modema_\alpha(X(t))$ information feature. As shown in Figure 2c,f, the $modema_\alpha()$ function greatly conditions the image of the sensing element's signal for both the resistance and the current as raw data information. The lower the α attenuation coefficient, the higher the dispersion of the $modema_\alpha(X(t))$ values over the whole acquisition. It is noticed that the information feature $modema_\alpha(X(t))$ is centered around zero. This property is a rather interesting feature if sensing elements' raw signal current or resistance displays large drifts over time, due to various physical mechanisms that may not necessarily threaten the quality of the information. It is also noticed that for any value of the α attenuation coefficient between $1/2$ and $1/1000$, the noise of the $modema_\alpha(X(t))$ is relatively comparable to the one present in the raw data. As the $ema_\alpha()$ function only filters the high-frequency signal of the floating reference, most of the high-frequency information associated with the raw data $X(t)$ is preserved in the $modema_\alpha(X(t))$ information descriptor. As this noise seems less sensitive to α, a large dependency of the $modema_\alpha(X(t))$ classification on α is expected, depending on whether the class-depending information is localized at the level of the information descriptor noise or its amplitude modulation. This has been verified by the unsupervised classification by PCA, focused strictly on the first two principal components (see Supplementary Materials for details on all α cases). Figure 2g,j show that the main information feature is not environmentally specific when using the raw data resistance or current as an information descriptor, with a temporal effect on the scores displayed in (PC1;PC2). Figure 2h,k show that using the floating reference $ema_\alpha(X(t))$ reduces greatly the score dispersion in the (PC1;PC2) projection for low values of α (we evidenced that the lower the filtering, the closer the PCA with the ema appears to the one with the raw data). However, using the floating reference $ema_\alpha(X(t))$ still does not allow clustering of the scores by the nature of the environment. The case of using the $modema_\alpha(X(t))$ as an information descriptor is substantially different from the one of the raw data $X(t)$ and the floating reference $ema_\alpha(X(t))$, for both the resistance and the current data as raw data values. One can observe in Figure 2i,l that at low values of α (effect of α detailed in the previous section), the information of the PCA score projection on (PC1;PC2) is organized by the nature of the environment class that is presented to the sensing hardware. PC1 specifically discriminates purge sequences from some acetone exposures, while PC2 specifically discriminates water exposures from ethanol ones, both when using the resistance and the current as raw data values. It is noticed that the dispersion of the scores is very comparable in both cases of using the resistance and the current as raw data, despite the results being different. This stresses the fact that the approach can be applied to resistance or current values without observing significant differences in the quality of the recognition; both classifications are distinct and may require a closer comparison of the explained variance for different values of the α attenuation coefficient in both cases (see Figure 3).

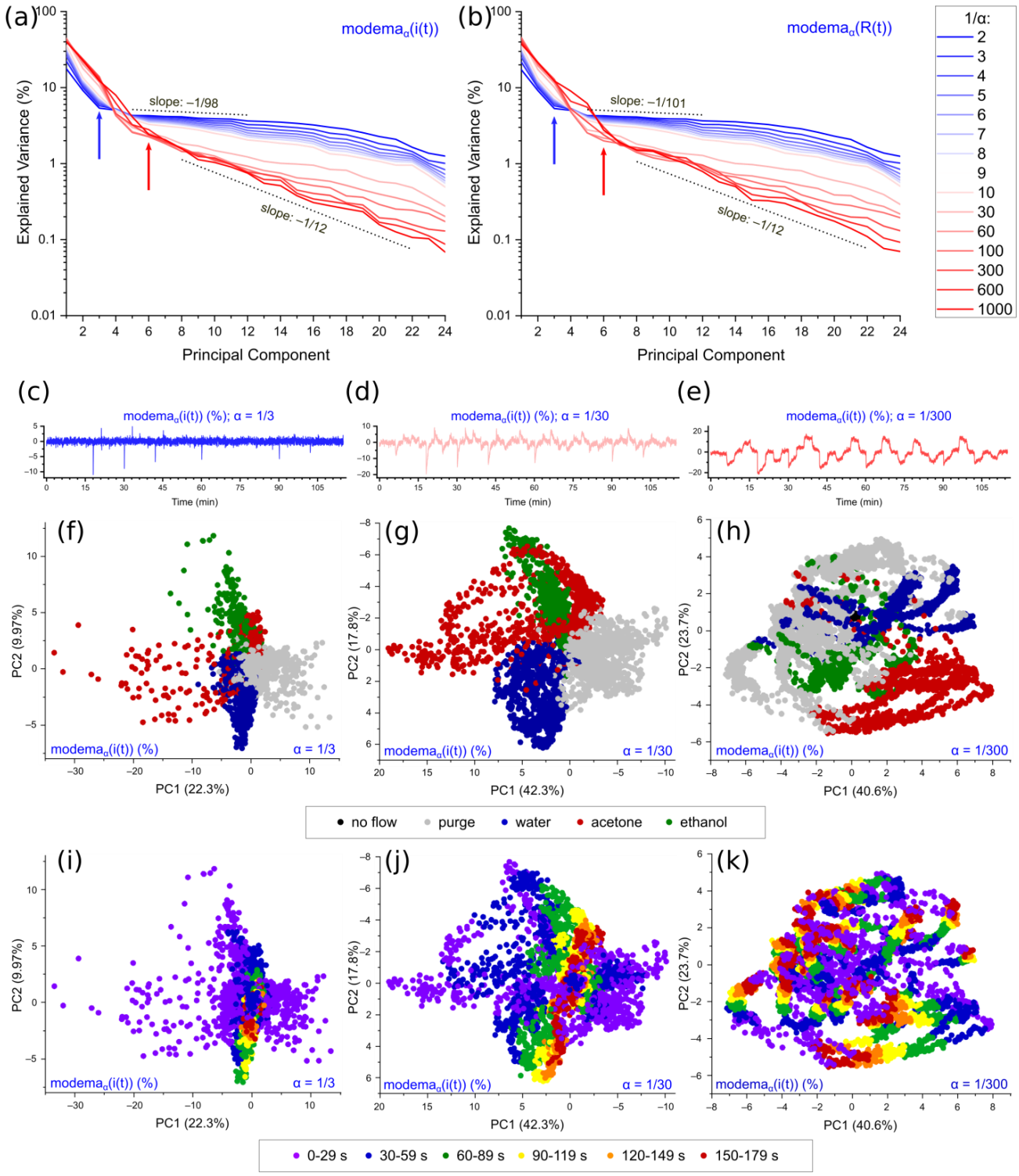

Figure 3. Data separation of the modema() information descriptors depending on the attenuation coefficient α. (**a**,**b**) Scree plots for $modema_\alpha(i(t))$ and $modema_\alpha(R(t))$ with different values of α. (**c**–**k**) First two component scores of the principal component analysis on the $modema_\alpha(i(t))$ datasets for $1/\alpha = 3$ (**c**,**f**,**k**), 30 (**d**,**g**,**j**) and 300 (**e**,**h**,**k**). (**c**–**e**) A single raw current data trace from Figure 2b for different values of $1/\alpha = 3$ (**c**), 30 (**d**) and 300 (**e**). (**f**–**k**) (PC1;PC2) planes for the full $modema_\alpha(i(t))$ dataset for $1/\alpha = 3$ (**f**,**i**), 30 (**g**,**j**) and 300 (**h**,**k**), labeled with the five environmental classes (**f**,**g**,**h**) or with six classes of acquisition times after exposure starting at t_0 (**i**,**j**,**k**).

3.3. Relationship between Attenuation Coefficient and Environment Clustering

The scree plots of the PCA analysis using $modema_\alpha(i(t))$ (Figure 3a) and $modema_\alpha(R(t))$ (Figure 3b) as distinct information descriptors quantify how much information variance is preserved in the different principal components of the scores. In both cases, very similar effects of the attenuation coefficient α are evidenced. The cumulative variance increases in PC1 and PC2 with the diminishing of the attenuation coefficient α (see Figure 3a,b). As the decrease in α is associated with an increase in the data dispersion in the information descriptor (see Figure 3c–e for $\alpha = 1/3$, $1/30$ and $1/300$, respectively), more information is contained in the first principal components by the choice of an attenuation coefficient that characterizes the modulations of signals upon exposures to upcoming classes.

This is in line with the idea that higher values of α lead to lower signal-to-noise ratios (where only fast transients are detected a few seconds after the change in an exposure), with most of the information being yielded by stochastic processes responsible for the preponderant noise observed in $modema_\alpha(i(t))$ and $modema_\alpha(R(t))$. The variance dispersion over the different principal components shows the same dependency with the attenuation coefficient α in both cases of using the $modema_\alpha(i(t))$ and $modema_\alpha(R(t))$ information features (see Figure 3a,b). By the elbow method (arrows displayed in Figure 3a,b), one can observe that most of the explained variance is gathered within the first three principal components in the case of a high attenuation coefficient, while in the case of applying a low attenuation coefficient, the explained variance is spread over the six principal components. This shows that PC3′s explained variance may promote significantly higher data separation by class of exposures in a three-dimensional graph (an example of the projection of the first three components for $1/\alpha = 300$ is available as Supplementary Materials in Figure S35). As the explained variance increases in higher principal components for the highest attenuation coefficients, this shows that when the $ema_\alpha()$ filter cuts off most of the transient signal, the information density is sparser (more noisy) but the remaining explainable signal is simpler (less parametric on a linear model). Instead, if the $ema_\alpha()$ only filters noise, more data can be extracted over the whole acquisition by the reading of the modulations in $modema_\alpha(X(t))$, but the information is harder to classify with only two principal components. Such a property shows that different data dispersion is to be expected, depending on the attenuation coefficient α applied on the $ema_\alpha()$ filter conditioning the $modema_\alpha(X(t))$ information descriptors. The following results specifically focus on using currents $X(t)$ as we expect similar classification dependencies with α in the case of using $R(t)$ as raw data information.

Principal component scores show various organizations on the (PC1;PC2) projection depending on the attenuation coefficient α (see Figure 3f–k). By looking at the gradual decrease in the attenuation coefficient α, it can be observed that data separability by the nature of the exposed environment tends to increase (comparison between $\alpha = 1/3$ and $1/30$ in Figure 3f,g), before the score clusters start collapsing in (PC1;PC2) (comparison between $\alpha = 1/30$ and $1/300$ in Figure 3g,h). This indicates that the $ema_\alpha(i(t))$ shall feature a floating reference point that is filtered from the inherent signal noise but not filtered from the whole activity of the signal. Such a characteristic presents similarities of both resistance modulation ($R/R_0 - 1$) and drift resistance ($\dot{R}/R$) as information descriptors, which gather different aspects of the signal's physics [30]. By the structure of the feature, both $modema_\alpha(i(t))$ and R/R_0 normalize a value recorded at a given time to a buffered reference that changes all along the acquisition. However, both $modema_\alpha(i(t))$ and $\dot{R}/R$ are referenced to a local value all along the acquisition. $modema_\alpha(X(t))$ can represent a generic compromise to reference data without the intervention of an external user (unlike setting $R/R_0 - 1$), with a higher representation of the steady effect of the environment on a material (unlike $\dot{R}/R$). When looking at the temporality of the data, it appears that the most separable feature samples are generically originating from the most dynamical raw data ones (as property shared with $\dot{R}/R$) [30]. Particularly for the cases of a high attenuation coefficient α, data generated within the first 30 s after a new class exposure at t_0 exhibit higher spreading on a (PC1;PC2) projection plane when compared to points recorded much

later for α = 1/3 and 1/30 (see Figure 3i,j). In these cases, the $ema_\alpha(i(t))$ floating reference is very similar to the raw signal $i(t)$. So, the $modema_\alpha(i(t))$ is very noisy except when transient (see Figure 3c,d). For a lower attenuation coefficient α at 1/300, it is also noticed that PCA scores on (PC1;PC2) are not organized by acquisition times anymore, as the $ema_\alpha(i(t))$ floating reference embeds less class-specific information and $modema_\alpha(i(t))$ reflects more a drift-free relative current modulation (see Figure 3e,k). It is particularly interesting to observe that, despite the transduction mechanism being interpreted as a thermodynamic property [28,50], resistance modulation is not the main source of information that an unsupervised classifier such as the PCA exploits to sort output current-related data by the nature of the molecular environment that is exposed on the doped polymers. As the quality of the PCA clustering is sensitive to the attenuation coefficient α, it may be important to inspect the PC loadings for different α to observe whether the classifier universally uses the same selection of materials, independently from what physical properties they may contribute to the $ema_\alpha(i(t))$ floating reference (see Figure 4).

3.4. Conducting Polymer Doping Complementarity in the Principal Component Analysis

As expected, loadings of PC1 and PC2 are very sensitive to the attenuation coefficient α, depending on whether $ema_\alpha()$ filters mostly the signal noise or all class-specific dynamics in the floating reference point (see Figure 4). The correlation circles show that doped polymer materials contribute differently in the principal components PC1 and PC2, with the attenuation coefficient α = 1/3, 1/30 or 1/300 (see Figure 4a–c). Particularly, pristine P3HT is distinctively the only material that contributes negatively to the first principal component PC1 for a lower α (see Figure 4a,b), while for a higher attenuation coefficient α, the material loses this property and contributes positively to PC1 like all the doped polymers (see Figure 4c). Surprisingly, one can observe that the direction of all loading vectors follows almost the same ordering in the correlation circles, and this is regardless of the value of the attenuation coefficient α. Particularly, it seems that such an order follows the same series of doping strengths that was previously defined by the nature of the dopant [28]. So, a stronger dopant, inducing higher values of the raw $i(t)$ data (see Figure 1e), seems to lead to a lower value for the phase coordinate on the loading vector of a sensing element in a correlation circle, while a moderately doped polymer appears to induce a higher phase coordinate for its sensing element loading vector. Here, one shall not directly attribute such a correlation to the causality of a generic doping mechanism in a conducting polymer for sensing in a machine learning framework. The correlation made between a physical property of the material, inducing a trend of data organization in PCA, is mostly induced by the fact that doping increases the nominal value of the $i(t)$ raw information, which indirectly increases the signal-to-noise ratio of the data collection. It shall, however, be highlighted that the higher sensitivity of $Ce(OTf)_3$ to water (observed in Figure 1e) greatly conditions the direction of the vector for its sensing elements in the different correlation circles, as PC2 allows for separating water samples from others. Regarding the absolute value of the PC loadings in PC1 and PC2, all the sensing elements have different contributions depending on the value of the attenuation coefficient α (see Figure 4d,e). Moreover, the inversion of contributions with α in the pristine P3HT loadings for PC1 is confirmed in Figure 3d, such as for one $Bi(OTf)_3$-sensing element and one $Dy(OTf)_3$ element over 15 different values of α. Furthermore, Figure 4e shows that an inversion of contributions with α in the PC2 loadings occurs for 12 different sensing elements over the 24 used in the experiment. It is observed that most contributions in the PC loadings at a given α are clearly material-dependent (see Figure 4d for $Fe(OTf)_3$-doped P3HT in particular and Figure 4e for $Ce(OTf)_3$-doped and pristine P3HT in particular). This shows that these contribution inversions induced by the α are no stochastic artifact and reflect a genuine property of the material dynamics. Therefore, it seems that using an unsupervised PCA linear classifier feels inadequate for generically ranking the quality of a material population using the ema floating reference. To quantify recognition at an optimal α, supervised learning was performed to evaluate the potential of modema for classification (see Figure 5).

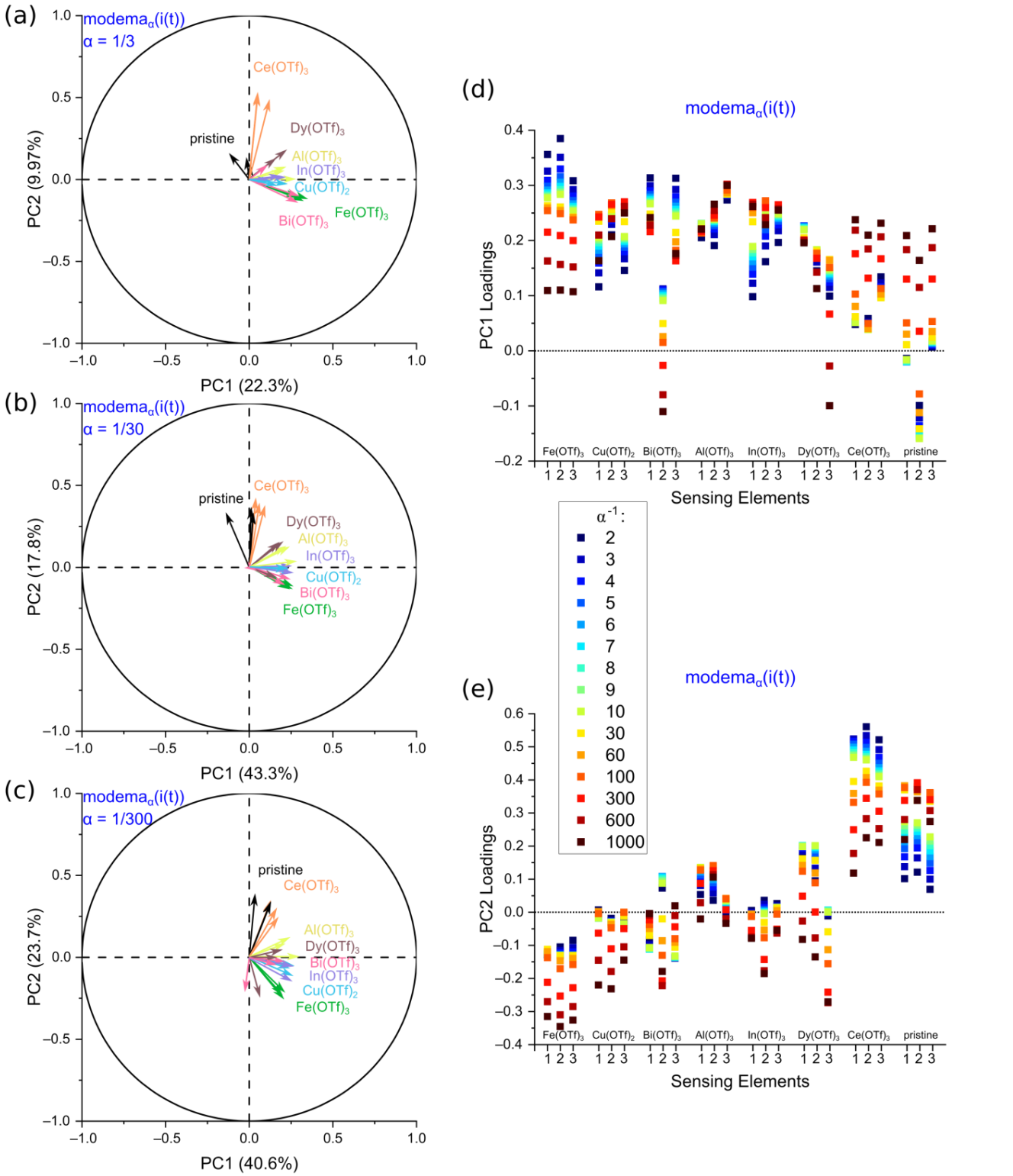

Figure 4. Contribution of the different materials in the PC1 and PC2 loadings. (**a**–**c**) Correlation circles in (PC1;PC2) for the *modema*$_\alpha$(*i*(*t*)) dataset of 24 sensing elements composing eight different metal triflate-doped and undoped P3HT coatings, for $1/\alpha = 3$ (**a**), 30 (**b**) and 300 (**c**). (**d**,**e**) Dependency of the α attenuation coefficient on the PC1 (**d**) and PC2 (**e**) loadings for the 24 sensing elements composed of the eight different metal triflate-doped and undoped P3HT coatings, for the PCA on the *modema*$_\alpha$(*i*(*t*)) feature as information descriptor.

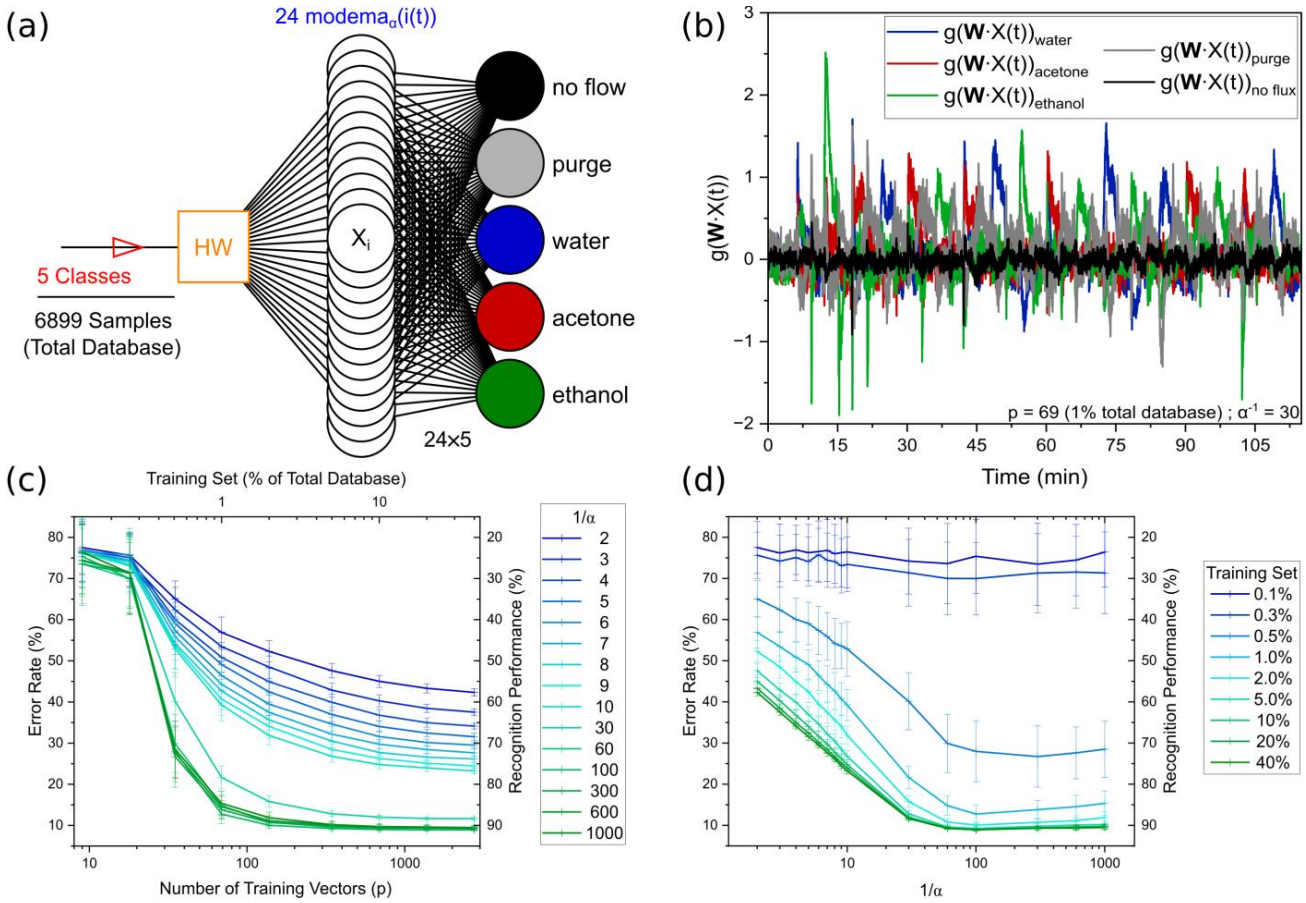

Figure 5. Environment recognition with the *modema*$_\alpha$(*i*(*t*)) information descriptor under supervised classification. (**a**) Software classifier structure used in this study for supervised classification of the five different classes of exposure, using the Moore–Penrose pseudo-inverse. (**b**) Example of the five different coordinates of $g(\mathbf{W}\cdot X(t))$ output vector for a single training, using $1/\alpha = 30$ and $p = 69$ vectors. (**c**) Dependency of the error rate on the five-class recognition with the size of the training set, for different values of the α attenuation coefficient. (**d**) Dependency of the error rate on the five-class recognition with the α attenuation coefficient, for different sizes of the training set.

3.5. Supervised Training of Environment Recognition with Output Currents' Modema

To quantify the performances of a classifier when recognizing molecular environments, the Moore–Penrose pseudo-inverse was used to supervise the training of the electronic nose in recognizing the five classes of environments by the use of *modema*$_\alpha$(*i*(*t*)) as the 24-dimensional information descriptor. In this case, the recognition space is not dimensionally reduced to only two principal components, so the *modema*$_\alpha$(*i*(*t*)) information is used in its entire complexity to recognize the feature vectors, whether they were recorded upon no flow, under purging or exposure of either a water, an acetone or an ethanol vapor blow (see Figure 5a). As in former studies [49], and similarly as in the case of the PCA classifier, all dynamical data generated during a complete acquisition are used in a classifier. Training and testing vectors are randomly picked from the 6899 available vectors to form two complementary databases (no validation database was used to optimize the attenuation coefficient α, as all the 15 α cases were generated in separated experiments as a variable of study). As a second variable of study, the size of the training dataset (related to the number of training vectors p) was studied to observe its impact on the classifier recognition performances at a given attenuation coefficient α. The classification threshold was set on the weighted $g(\mathbf{W}\cdot X(t))$ vectors (with *modema*$_\alpha$(*i*(*t*)) as $X(t)$) optimized by pseudo-inversion after training, showing the different components of the five environments to express mostly at different periods of time during the acquisition (see Figure 5b).

For different training/testing cases, it was observed that the quality of the classification was highly dependent on both variables of study: the attenuation coefficient α of the 24-dimensional $moderma_{\alpha}(i(t))$ input and the number of training vectors p used for training/calibrating the classifier (see Figure 5c,d). Naturally, the larger the dataset, the lower the error rate at a given α. Also, the lowest attenuation coefficient α led systematically to the highest recognition rate at a given size for a training dataset. Both properties were statistically verified, as each case displayed in Figure 5c,d represents the average value of the error rate for 100 cases of randomly picked training vectors in each scenario of fixed α and p parameters. The trend with p is comparable to the former study on binary classification, using the same electrode structures as an OECT array to recognize projected gate voltage patterns [49]. However, it is noticed that a minimum of 0.5% of the database is required to significantly train the system to recognize the five classes, from 75 $\pm$ 2% error (quasi independently from α) down to 25 $\pm$ 15% when trained with 5% of the database or higher (see Figure 5c). The impact of the attenuation coefficient α is also quite significant on the error when the classifier is trained enough (see Figure 5d). In the case where at least 0.5% of the database is used to train the classifier, a transition at $\alpha = 1/30$ is observed: for lower α values, the system does not progress further at a given p, and for higher α values, the classifier progresses at a quasi-linear rate with $log(1/\alpha)$, as displayed in Figure 5d. This result is quite surprising as it shows, contrary to the unsupervised classification by PCA, that the classification performs better under supervised training when the attenuation coefficient α is minimized.

3.6. Conducting Polymer Doping Complementarity with the Moore–Penrose Pseudo-Inverse

It is therefore particularly important to stress that the figures of merit for sensing array material selection (see Figure 6) are highly dependent on the machine learning approach for a given application. The fact that the supervised linear classifier favors the use of a highly attenuated $ema_{\alpha}(i(t))$ floating reference, as opposed to the unsupervised PCA classifier that favors the least attenuated $ema_{\alpha}(i(t))$, motivates us to further understand whether both classifiers exploit the same sensing element contributions in their weights/loadings or not (see Figure 6). A summary of Figure 5c,d is represented (without statistics) as a correlation heatmap between p and $1/\alpha$ in Figure 6a. Using the linear color gradients with the error rate, an optimal case between training dataset sparsity and reference signal attenuation is identified for $p = 69$ (1% total database) and $\alpha = 1/30$ to ensure 21.6% recognition on average. As the supervised approach requires large statistics of train-and-tests compared to PCA, the material selection optimization has been assessed only for the specific value of $\alpha = 1/30$ for supervised classification, so no weight inversions with the attenuation coefficient are verified in the following. For the specific case of $(p;\alpha) = (69;1/30)$, the confusion matrix displays the statistics of true/false positives/negatives of 1000 train-and-tests (see Figure 6b). During the acquisition, "purge" samples occur the most, while "no flow" ones are less common than each of the water, acetone and ethanol vapor exposures. Therefore, predictive rates are naturally higher in recognizing "purge" (91%) than "no flow" (16%), with recognition of the three different vapors between 71 and 76% (see Figure 6b). However, true positive rates are particularly in favor of classifying the three different vapors at 86–87% recognition rates, compared to "purges" (73%) and "no flow" (23%). It indicates that the classifier is particularly sensitive to exposures of vapors compared to both "no flow" and "purge" carrying no solvent vapors: upon exposure to "acetone", "ethanol" or "water", environments are recognized in 85–86% of the cases with as little as 1% of the database for training, with $ema_{\alpha}(i(t))$ if $\alpha = 1/30$. The material selection by the efficient supervised classifier at $(p;\alpha) = (69;1/30)$ is a bit less difficult to interpret than the latter case of the PCA classification (see Figure 6c). It is observed that no material is of particularly generic relevance for identifying "no flow" (previously, this class was not separable from the rest in the (PC1;PC2) projection). For the other classes, it seems $Bi(OTf)_3$-doped P3HT (a highly doped material) was particularly useful in inducing negative weights, but not in a reproducible way between the three sensing devices. A similar statement can be used for

pristine P3HT (a low-doping material), which induces the highest weights on the "acetone" output class. In general, the standard deviation of each weight is quite significant compared to the actual value of the weight. In conclusion, it still seems quite difficult to infer the material-set quality to efficiently recognize the vapors under supervised learning with a linear classifier, despite its high sensitivity to specifically recognize "water", "acetone" and "ethanol" with the ema floating reference in the $modema_{\alpha}(i(t))$ information descriptor.

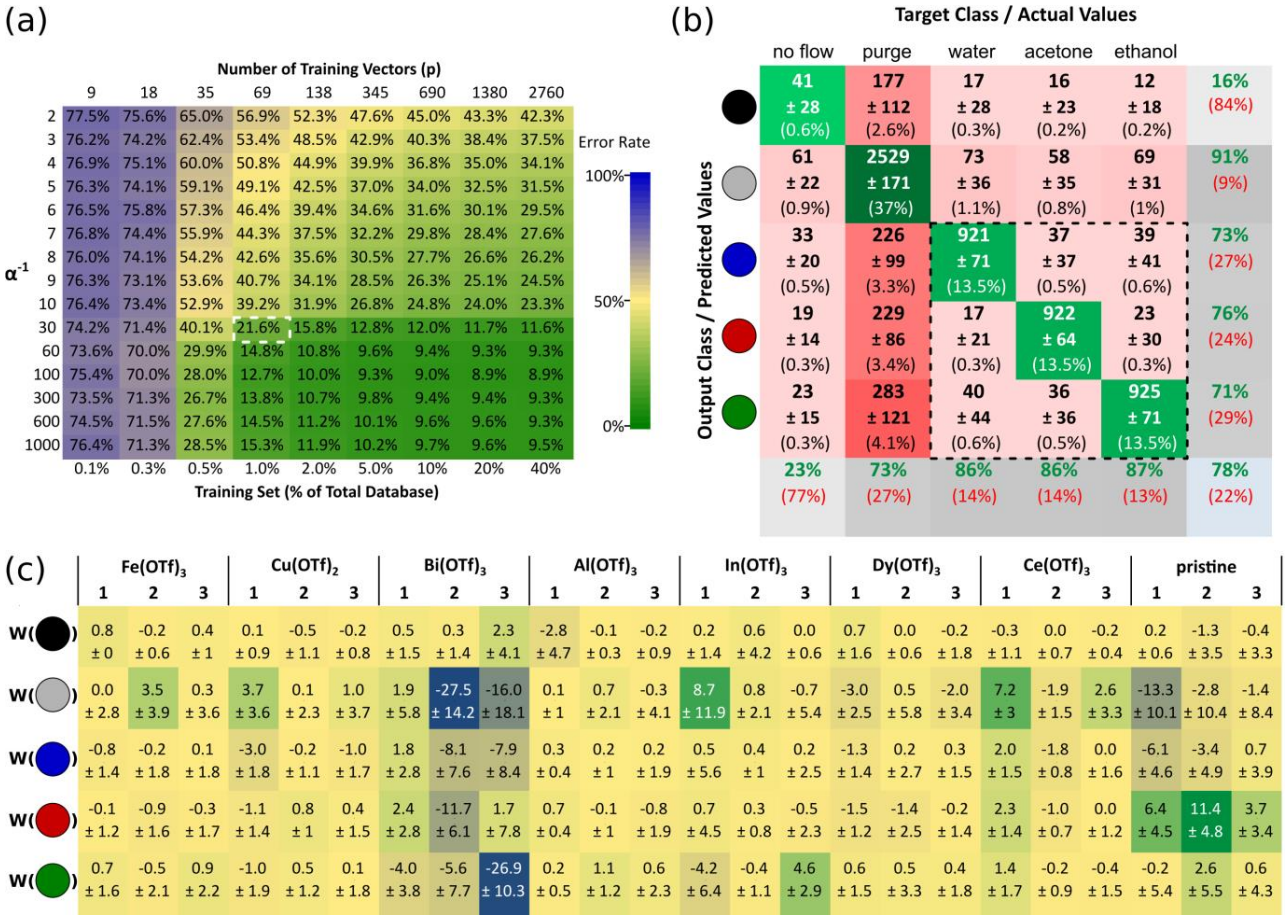

Figure 6. Material/environment specificity in the case of supervised classification with the $modema_{\alpha}(i(t))$ information descriptor. (**a**) Correlation heatmap between the attenuation coefficient α and the size of the training dataset for the recognition of the five different classes of environments, using the supervised classifier. Each displayed error rate is calculated as the mean value of 100 training/testing sets with given α and p values of different randomly selected sets of training vectors. (**b**) Confusion matrix for the supervised classification of the five environments. The displayed values are calculated as the mean and standard deviation values of 1000 training/testing sets with $1/\alpha = 30$ and $p = 69$ with different randomly selected sets of training vectors. (**c**) Correlation heatmap between the weights connecting the different environmental classes to the 24 sensing elements composed of eight different metal triflate-doped and undoped P3HT coatings, after supervised classification. The displayed values are calculated as the mean and standard deviation values of 1000 training/testing sets with $1/\alpha = 30$ and $p = 69$ with different randomly selected sets of training vectors.

4. Discussion

Resistance or current modulations as classical information descriptors for conductimetric electronic noses are experiment-dependent parameters that require fixing references periodically at given times during the exposures and being synchronized with the sample exposures. As such, they can hardly be considered generic figures of merit in conductimetric electronic noses. Using $ema_{\alpha}()$ referencing may be better suited in this sense as it

does not require external parameterizing, except setting the attenuation coefficient α. Such a value can also be data-dependent to vary dynamically and be generic in absolute. In this study, the analysis, both under unsupervised and supervised learning, shows high sensitivity of the classifier with α. Thus, the choice for data-dependent functions that adjust α over time, depending on the dynamic dispersion of a high-dimensional signal at a given time, is of the highest interest to define a figure of merit for conductimetric electronic noses. In the case of doped conducting polymers, PCA as an unsupervised classifier uses all doped P3HT contributions to recognize the transient points with the least attenuated floating reference. It also appeared that the doping strengths of the different metal triflates have a crucial role in their different contributions to the first two principal components. The pseudo-inverse supervised classifier instead performs better with a highly attenuated floating reference. However, the involvement of the different materials in the analysis is extremely sensitive to the training database. It is very surprising to observe that for the same collection of raw data generated from the same sensing array classifying the same environment classes, two different linear classifiers may exploit very different information for recognition, as the classifiers perform better with different attenuation coefficients. It is therefore with the highest care that we need to consider $modema_{\alpha}(i(t))$ as a figure of merit in a specific classification framework, in case α is fixed. Although it may need to be confirmed with other materials, the quality of the classification at a fixed α is highly dependent on the learning scheme: supervised or not supervised.

Generating relevant information, either near-sensor or in-sensor, directly from each physical device/sensing node is of major importance to lower the computational complexity of a software classifier for inferring instantly [51,52]. Furthermore, in materio information preprocessing is an essential key to lowering this complexity by the use of materials that specifically convolute signals thanks to inner physical mechanisms but also to lower the fabrication costs for such classifiers [53,54]. In this sense, using $modema_{\alpha}(i(t))$ represents an extraordinary opportunity to revise the way we build information generators in electronics, in a neuro-inspired way, to emulate it closer to the way an actual receptor neuron conditions classification.

On this matter, information in the olfactory cortex in a human cerebrum is generated from about 6×10^6 olfactory receptor neurons gathering 339 different functional G-coupled receptors, which project the information in parallel [55,56]. Information carriers are particularly slow for olfaction (second-scale), as metabotropic receptors are generally slower than ionotropic ones but also because of the diffusion and retention times of volatile organic compounds in the mucus [57,58]. Very similarly to a conductimetric nose, a biological nose may use both the chemospecificity of metabotropic receptors on each cell to recognize odors but also the kinetics of their interactions with molecules. Also, the brain itself uses information temporality in a relevant way for classification. Many neuron models have attempted to conceptualize the transience of electrical responses responsible for the neurons' plasticity as a passive electric circuit. Most of them associate the output response of a neuron with the charge of a serial capacitor, such as Hodgkin and Huxley's model or the leaky-integrate-and-fire model derived from Lapicque's earlier works [59]. Unconventional computing technologies mimicking such plasticity, which embedding RC elements in their Voigt model expression [60], can use the temporality of the information signal to classify sensed information with doped conducting polymers [61]. Inherent electrochemical processes associated with these same doped polymers can even exhibit non-ideal capacitances [62,63], for which the analytical expression of their admittance in the time domain shows true biological and computational relevance to tuning the synaptic temporality [64,65]. The capability to tune such non-ideality in sensing devices may be a true hardware feature for sorting the information in materio by tuning the fading-memory time window of the sensing element, as a physical model of a biological receptor neuron.

In this study, a strong parallel may be identified between the influence of the attenuation coefficient α in the ema floating reference function on the quality of environment recognition and the influence of the non-ideality factor α of a fractional leaky-integrate-and-

fire (FLIF) neuron [64,65]. In particular, this study spots the fact that controlling the ability to tune the attenuation coefficient α may condition the $modema_{\alpha}(i(t))$ feature output to use different sensitivities from the materials depending on the learning framework. It has to be stressed at this point that the analytical expression of the $modema_{\alpha}(i(t))$ information feature is not comparable to the electrical output of a FLIF neuron, the same way the attenuation coefficient of the ema function is not to be identified as the fractional order of the derivation in the FLIF neuron expression. However, the fact that doped conducting polymers are both able to classify environments depending on their attenuation coefficient and able to use inherent electrochemical processes to change their capacitive properties represents a very encouraging perspective for in materio information classification within the same class of material, such as doped conducting polymers in sensing hardware.

5. Conclusions

This study focuses on the "modema" information descriptor as a potential figure of merit for conductimetric electronic noses. By the use of several doped conducting polymers, the ema floating point has shown to be a good reference for extracting relevant information from a whole acquisition to classify volatile molecular environments. Under supervised learning, a linear classifier is able to recognize 90% of five classes, regardless of the dynamic of the data. Under unsupervised learning, a linear classifier can cluster data by environment identity, whether an air blow is loaded with water, ethanol or acetone vapors or not. The great advantage of such a feature is that it does not need any user- nor environment-specific post-parameterization, which makes its application quite generic: for analyzing volatile samples intermittently for product quality monitoring, or for online analysis of time-varying environment patterns such as outdoor pollution. The sensitivity of the attenuation coefficient to the learning framework (either supervised or not supervised) is both an important point to optimize prior to using such a feature for environment pattern classification and an interesting property of the descriptor to investigate, for identifying "good" material combinations for a sensing input depending on the structure of the data. Further perspectives are to be envisioned both at software and hardware levels: to optimize the descriptor so the attenuation coefficient can self-adapt to the data organization and to emulate such filtering in materio or near the sensors at the input of an electronic nose, to better mimic a biological sense in the way environmental information is sorted.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/electronics13030497/s1, Figure S1: PCA on "$modema_{\alpha}(i(t))$" for $\alpha = 1/2$; Figure S2: PCA on "$modema_{\alpha}(i(t))$" for $\alpha = 1/3$; Figure S3: PCA on "$modema_{\alpha}(i(t))$" for $\alpha = 1/4$; Figure S4: PCA on "$modema_{\alpha}(i(t))$" for $\alpha = 1/5$; Figure S5: PCA on "$modema_{\alpha}(i(t))$" for $\alpha = 1/6$; Figure S6: PCA on "$modema_{\alpha}(i(t))$" for $\alpha = 1/7$; Figure S7: PCA on "$modema_{\alpha}(i(t))$" for $\alpha = 1/8$; Figure S8: PCA on "$modema_{\alpha}(i(t))$" for $\alpha = 1/9$; Figure S9: PCA on "$modema_{\alpha}(i(t))$" for $\alpha = 1/10$; Figure S10: PCA on "$modema_{\alpha}(i(t))$" for $\alpha = 1/30$; Figure S11: PCA on "$modema_{\alpha}(i(t))$" for $\alpha = 1/60$; Figure S12: PCA on "$modema_{\alpha}(i(t))$" for $\alpha = 1/100$; Figure S13: PCA on "$modema_{\alpha}(i(t))$" for $\alpha = 1/300$; Figure S14: PCA on "$modema_{\alpha}(i(t))$" for $\alpha = 1/600$; Figure S15: PCA on "$modema_{\alpha}(i(t))$" for $\alpha = 1/1000$; Figure S16: PCA on "$modema_{\alpha}(R(t))$" for $\alpha = 1/2$; Figure S17: PCA on "$modema_{\alpha}(R(t))$" for $\alpha = 1/3$; Figure S18: PCA on "$modema_{\alpha}(R(t))$" for $\alpha = 1/4$; Figure S19: PCA on "$modema_{\alpha}(R(t))$" for $\alpha = 1/5$; Figure S20: PCA on "$modema_{\alpha}(R(t))$" for $\alpha = 1/6$; Figure S21: PCA on "$modema_{\alpha}(R(t))$" for $\alpha = 1/7$; Figure S22: PCA on "$modema_{\alpha}(R(t))$" for $\alpha = 1/8$; Figure S23: PCA on "$modema_{\alpha}(R(t))$" for $\alpha = 1/9$; Figure S24: PCA on "$modema_{\alpha}(R(t))$" for $\alpha = 1/10$; Figure S25: PCA on "$modema_{\alpha}(R(t))$" for $\alpha = 1/30$; Figure S26: PCA on "$modema_{\alpha}(R(t))$" for $\alpha = 1/60$; Figure S27: PCA on "$modema_{\alpha}(R(t))$" for $\alpha = 1/100$; Figure S28: PCA on "$modema_{\alpha}(R(t))$" for $\alpha = 1/300$; Figure S29: PCA on "$modema_{\alpha}(R(t))$" for $\alpha = 1/600$; Figure S30: PCA on "$modema_{\alpha}(R(t))$" for $\alpha = 1/1000$; Figure S31: PCA on "$ema_{\alpha}(i(t))$" for $\alpha = 1/2$; Figure S32: PCA on "$ema_{\alpha}(i(t))$" for $\alpha = 1/1000$; Figure S33: PCA on "$ema_{\alpha}(R(t))$" for $\alpha = 1/2$; Figure S34: PCA on "$ema_{\alpha}(R(t))$" for $\alpha = 1/1000$; Figure S35: PCA on "$modema_{\alpha}(R(t))$" for $\alpha = 1/2$.

Author Contributions: Conceptualization, S.P.; Methodology, S.P.; Validation, W.H.A.; Formal analysis, W.H.A. and A.B. (Aicha Boujnah); Investigation, W.H.A., A.B. (Aicha Boujnah), A.B. (Antoine Baron) and S.P.; Writing—original draft, W.H.A., A.B. (Antoine Baron) and S.P.; Writing—review & editing, A.B. (Aicha Boujnah), A.B. (Aimen Boubaker), A.K. and K.L.; Visualization, S.P.; Supervision, S.P.; Project administration, A.B. (Aimen Boubaker), A.K. and S.P.; Funding acquisition, A.B. (Aimen Boubaker), A.K. and S.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the French National Research Agency (ANR), grant number ANR-22-CE24-0001-01. The APC was funded by the editor.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Materials; further inquiries can be directed to the corresponding author.

Acknowledgments: The authors thank the French Nanofabrication Network RENATECH for financial support for the IEMN cleanroom. The authors appreciate the inspiring work of Alexander Vergara on machine olfaction.

Conflicts of Interest: The authors declare no conflicts interests.

References

1. Ullah, A.; Anwar, S.M.; Li, J.; Nadeem, L.; Mahmood, T.; Rehman, A.; Saba, T. Smart cities: The role of Internet of Things and machine learning in realizing a data-centric smart environment. *Complex Intell. Syst.* **2023**, 1–31. [CrossRef]
2. Liu, Z.; Xiao, G.; Liu, H.; Wei, H. Multi-sensor measurement and data fusion. *IEEE Instrum. Meas. Mag.* **2022**, *25*, 28–36. [CrossRef]
3. Schiavi, A.; Mazzoleni, F.; Facello, A.; Prato, A. Metrology for next generation "Phygital Sensors". In Proceedings of the 2023 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4. 0&IoT), Brescia, Italy, 6–8 June 2023; IEEE: Piscataway, NJ, USA, 2023.
4. Kim, C.; Lee, K.K.; Kang, M.S.; Shin, D.-M.; Oh, J.-W.; Lee, C.-S.; Han, D.-W. Artificial olfactory sensor technology that mimics the olfactory mechanism: A comprehensive review. *Biomater. Res.* **2022**, *26*, 40. [CrossRef]
5. Wise, P.M.; Olsson, M.J.; Cain, W.S. Quantification of odor quality. *Chem. Senses* **2000**, *25*, 429–443. [CrossRef] [PubMed]
6. Bushdid, C.; Magnasco, M.O.; Vosshall, L.B.; Keller, A. Humans can discriminate more than 1 trillion olfactory stimuli. *Science* **2014**, *343*, 1370–1372. [CrossRef] [PubMed]
7. Mamlouk, A.M.; Martinetz, T. On the dimensions of the olfactory perception space. *Neurocomputing* **2004**, *58*, 1019–1025. [CrossRef]
8. Koulakov, A.A.; Kolterman, B.E.; Enikolopov, A.G.; Rinberg, D. In search of the structure of human olfactory space. *Front. Syst. Neurosci.* **2011**, *5*, 65. [CrossRef]
9. Zhou, Y.; Smith, B.H.; Sharpee, T.O. Hyperbolic geometry of the olfactory space. *Sci. Adv.* **2018**, *4*, eaaq1458. [CrossRef]
10. Billesbølle, C.B.; de March, C.A.; van der Velden, W.J.C.; Ma, N.; Tewari, J.; del Torrent, C.L.; Li, L.; Faust, B.; Vaidehi, N.; Matsunami, H.; et al. Structural basis of odorant recognition by a human odorant receptor. *Nature* **2023**, *615*, 742–749. [CrossRef]
11. Castro, J.B.; Ramanathan, A.; Chennubhotla, C.S. Categorical Dimensions of Human Odor Descriptor Space Revealed by Non-Negative Matrix Factorization. *PLoS ONE* **2013**, *8*, e73289. [CrossRef]
12. Saini, K.; Ramanathan, V. Predicting odor from molecular structure: A multi-label classification approach. *Sci. Rep.* **2022**, *12*, 13863. [CrossRef] [PubMed]
13. Lee, B.K.; Mayhew, E.J.; Sanchez-Lengeling, B.; Wei, J.N.; Qian, W.W.; A Little, K.; Andres, M.; Nguyen, B.B.; Moloy, T.; Yasonik, J.; et al. A principal odor map unifies diverse tasks in olfactory perception. *Science* **2023**, *381*, 999–1006. [CrossRef] [PubMed]
14. Friedman, L.; Miller, J.G. Odor incongruity and chirality. *Science* **1971**, *172*, 1044–1046. [CrossRef] [PubMed]
15. Abate, A.; Brenna, E.; Fuganti, C.; Gatti, F.G.; Giovenzana, T.; Malpezzi, L.; Serra, S. Chirality and fragrance chemistry: Stereoisomers of the commercial chiral odorants Muguesia and Pamplefleur. *J. Org. Chem.* **2005**, *70*, 1281–1290. [CrossRef] [PubMed]
16. Bentley, R. The nose as a stereochemist. *Enantiomers and odor. Chem. Rev.* **2006**, *106*, 4099–4112. [CrossRef] [PubMed]
17. Gronenberg, W.; Raikhelkar, A.; Abshire, E.; Stevens, J.; Epstein, E.; Loyola, K.; Rauscher, M.; Buchmann, S. Honeybees (*Apis mellifera*) learn to discriminate the smell of organic compounds from their respective deuterated isotopomers. *Proc. R. Soc. B Biol. Sci.* **2014**, *281*, 20133089. [CrossRef]
18. Genva, M.; Kemene, T.K.; Deleu, M.; Lins, L.; Fauconnier, M.-L. Is It Possible to Predict the Odor of a Molecule on the Basis of its Structure? *Int. J. Mol. Sci.* **2019**, *20*, 3018. [CrossRef]
19. Verma, P.; Panda, S. Polymer selection approaches for designing electronic noses: A comparative study. *Sens. Actuators B Chem.* **2018**, *273*, 365–376. [CrossRef]
20. Park, S.Y.; Kim, Y.; Kim, T.; Eom, T.H.; Kim, S.Y.; Jang, H.W. Chemoresistive materials for electronic nose: Progress, perspectives, and challenges. *InfoMat* **2019**, *1*, 289–316. [CrossRef]
21. Rajagopalan, A.K.; Petit, C. Material Screening for Gas Sensing Using an Electronic Nose: Gas Sorption Thermodynamic and Kinetic Considerations. *ACS Sens.* **2021**, *6*, 3808–3821. [CrossRef]

22. Hatfield, J.; Neaves, P.; Hicks, P.; Persaud, K.; Travers, P. Towards an integrated electronic nose using conducting polymer sensors. *Sens. Actuators B Chem.* **1994**, *18*, 221–228. [CrossRef]
23. Freund, M.S.; Lewis, N.S. A chemically diverse conducting polymer-based "electronic nose". *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 2652–2656. [CrossRef] [PubMed]
24. Sierra-Padilla, A.; García-Guzmán, J.J.; López-Iglesias, D.; Palacios-Santander, J.M.; Cubillana-Aguilera, L. E-Tongues/Noses Based on Conducting Polymers and Composite Materials: Expanding the Possibilities in Complex Analytical Sensing. *Sensors* **2021**, *21*, 4976. [CrossRef] [PubMed]
25. Lüssem, B.; Riede, M.; Leo, K. Doping of organic semiconductors. *Phys. Status Solidi A* **2013**, *210*, 9–43. [CrossRef]
26. Salzmann, I.; Heimel, G.; Oehzelt, M.; Winkler, S.; Koch, N. Molecular Electrical Doping of Organic Semiconductors: Fundamental Mechanisms and Emerging Dopant Design Rules. *Acc. Chem. Res.* **2016**, *49*, 370–378. [CrossRef] [PubMed]
27. Sakai, N.; Warren, R.; Zhang, F.; Nayak, S.; Liu, J.; Kesava, S.V.; Lin, Y.-H.; Biswal, H.S.; Lin, X.; Grovenor, C.; et al. Adduct-based p-doping of organic semiconductors. *Nat. Mater.* **2021**, *20*, 1248–1254. [CrossRef]
28. Boujnah, A.; Boubaker, A.; Kalboussi, A.; Lmimouni, K.; Pecqueur, S. Mildly-doped polythiophene with triflates for molecular recognition. *Synth. Met.* **2021**, *280*, 116890. [CrossRef]
29. Boujnah, A.; Boubaker, A.; Pecqueur, S.; Lmimouni, K.; Kalboussi, A. An electronic nose using conductometric gas sensors based on P3HT doped with triflates for gas detection using computational techniques (PCA, LDA, and kNN). *J. Mater. Sci. Mater. Electron.* **2022**, *33*, 27132–27146. [CrossRef]
30. Ammar, W.H.; Boujnah, A.; Boubaker, A.; Kalboussi, A.; Lmimouni, K.; Pecqueur, S. Steady vs. Dynamic Contributions of Different Doped Conducting Polymers in the Principal Components of an Electronic Nose's Response. *Eng* **2023**, *4*, 2483–2496. [CrossRef]
31. Cipriano, D.; Capelli, L. Evolution of Electronic Noses from Research Objects to Engineered Environmental Odour Monitoring Systems: A Review of Standardization Approaches. *Biosensors* **2019**, *9*, 75. [CrossRef]
32. Capelli, L.; Sironi, S.; Del Rosso, R. Electronic Noses for Environmental Monitoring Applications. *Sensors* **2014**, *14*, 19979–20007. [CrossRef] [PubMed]
33. Wilson, A.D. Review of Electronic-nose technologies and algorithms to detect hazardous chemicals in the environment. *Procedia Technol.* **2012**, *1*, 453–463. [CrossRef]
34. Wang, M.; Chen, Y. Electronic nose and its application in the food industry: A review. *Eur. Food Res. Technol.* **2023**, *250*, 21–67. [CrossRef]
35. Tan, J.; Xu, J. Applications of electronic nose (e-nose) and electronic tongue (e-tongue) in food quality-related properties determination: A review. *Artif. Intell. Agric.* **2020**, *4*, 104–115. [CrossRef]
36. Shi, H.; Zhang, M.; Adhikari, B. Advances of electronic nose and its application in fresh foods: A review. *Crit. Rev. Food Sci. Nutr.* **2017**, *58*, 2700–2710. [CrossRef]
37. Cheng, L.; Meng, Q.-H.; Lilienthal, A.J.; Qi, P.-F. Development of compact electronic noses: A review. *Meas. Sci. Technol.* **2021**, *32*, 062002. [CrossRef]
38. Tiele, A.; Wicaksono, A.; Ayyala, S.K.; Covington, J.A. Development of a compact, IoT-enabled electronic nose for breath analysis. *Electronics* **2020**, *9*, 84. [CrossRef]
39. Jaeschke, C.; Padilla, M.; Turppa, E.; Polaka, I.; Gonzalez, O.; Richardson, K.; Pajukanta, J.; Kortelainen, J.M.; Shani, G.; Shuster, G.; et al. Overview on SNIFFPHONE: A portable device for disease diagnosis. In Proceedings of the 2019 IEEE International Symposium on Olfaction and Electronic Nose (ISOEN), Fukuoka, Japan, 26–29 May 2019; IEEE: Piscataway, NJ, USA, 2019.
40. Ye, Z.; Liu, Y.; Li, Q. Recent progress in smart electronic nose technologies enabled with machine learning methods. *Sensors* **2021**, *21*, 7620. [CrossRef]
41. Faleh, R.; Othman, M.; Gomri, S.; Aguir, K.; Kachouri, A. A transient signal extraction method of WO_3 gas sensors array to identify polluant gases. *IEEE Sens. J.* **2016**, *16*, 3123–3130. [CrossRef]
42. Yan, J.; Guo, X.; Duan, S.; Jia, P.; Wang, L.; Peng, C.; Zhang, S. Electronic nose feature extraction methods: A review. *Sensors* **2015**, *15*, 27804–27831. [CrossRef]
43. Muezzinoglu, M.K.; Vergara, A.; Huerta, R.; Rulkov, N.; Rabinovich, M.I.; Selverston, A.; Abarbanel, H.D.I. Acceleration of chemo-sensory information processing using transient features. *Sens. Actuators B Chem.* **2009**, *137*, 507–512. [CrossRef]
44. Vergara, A.; Vembu, S.; Ayhan, T.; Ryan, M.A.; Homer, M.L.; Huerta, R. Chemical gas sensor drift compensation using classifier ensembles. *Sens. Actuators B Chem.* **2012**, *166*, 320–329. [CrossRef]
45. John, A.D. Lange's handbook of chemistry. In *Universitas of Tennese Knoxville*, 15th ed.; Mc. Graw Hill Inc.: New York, NY, USA, 1999.
46. Lin, Z.; Cong, W.; Zhang, J.A. Biobutanol Production from Acetone–Butanol–Ethanol Fermentation: Developments and Prospects. *Fermentation* **2023**, *9*, 847. [CrossRef]
47. Zhang, X.; Guo, S.-X.; Gandionco, K.A.; Bond, A.M.; Zhang, J. Electrocatalytic carbon dioxide reduction: From fundamental principles to catalyst design. *Mater. Today Adv.* **2020**, *7*, 100074. [CrossRef]
48. Metsalu, T.; Vilo, J. ClustVis: A web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res.* **2015**, *43*, W566–W570. [CrossRef] [PubMed]
49. Pecqueur, S.; Talamo, M.M.; Guérin, D.; Blanchard, P.; Roncali, J.; Vuillaume, D.; Alibart, F. Neuromorphic time-dependent pattern classification with organic electrochemical transistor arrays. *Adv. Electron. Mater.* **2018**, *4*, 1800166. [CrossRef]

50. Ferchichi, K.; Bourguiga, R.; Lmimouni, K.; Pecqueur, S. Concentration-control in all-solution processed semiconducting polymer doping and high conductivity performances. *Synth. Met.* **2020**, *262*, 116352. [CrossRef]
51. Zhou, F.; Chai, Y. Near-sensor and in-sensor computing. *Nat. Electron.* **2020**, *3*, 664–671. [CrossRef]
52. Chai, Y.; Liao, F. Near-sensor and In-sensor Computing. *Nat. Electron.* **2022**, *3*, 664–671. [CrossRef]
53. Ghazal, M.; Mansour, M.D.; Scholaert, C.; Dargent, T.; Coffinier, Y.; Pecqueur, S.; Alibart, F. Bio-inspired adaptive sensing through electropolymerization of organic electrochemical transistors. *Adv. Electron. Mater.* **2022**, *8*, 2100891. [CrossRef]
54. Scholaert, C.; Janzakova, K.; Coffinier, Y.; Alibart, F.; Pecqueur, S. Plasticity of conducting polymer dendrites to bursts of voltage spikes in phosphate buffered saline. *Neuromorphic Comput. Eng.* **2022**, *2*, 044010. [CrossRef]
55. Moran, D.T.; Rowley, J.C.; Jafek, B.W.; Lovell, M.A. The fine structure of the olfactory mucosa in man. *J. Neurocytol.* **1982**, *11*, 721–746. [CrossRef] [PubMed]
56. Malnic, B.; Godfrey, P.A.; Buck, L.B. The human olfactory receptor gene family. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 2584–2589. [CrossRef] [PubMed]
57. Ghatpande, A.S.; Reisert, J. Olfactory receptor neuron responses coding for rapid odour sampling. *J. Physiol.* **2011**, *589*, 2261–2273. [CrossRef] [PubMed]
58. Purves, D.; Augustine, G.J.; Fitzpatrick, D.; Hall, W.; LaMantia, A.-S.; McNamar, J.O.; Williams, S.M. *Neurosciences*, 3rd ed.; Sinauer Associates Inc.: Sunderland, MA, USA, 2003.
59. Gerstner, W.; Kistler, W.M. *Spiking Neuron Models: Single Neurons, Populations, Plasticity*; Cambridge University Press: Cambridge, UK, 2002.
60. Pecqueur, S.; Lončarić, I.; Zlatić, V.; Vuillaume, D.; Crljen, Ž. The non-ideal organic electrochemical transistors impedance. *Org. Electron.* **2019**, *71*, 14–23. [CrossRef]
61. Pecqueur, S.; Vuillaume, D.; Crljen, Ž.; Lončarić, I.; Zlatić, V. A Neural Network to Decipher Organic Electrochemical Transistors' Multivariate Responses for Cation Recognition. *Electron. Mater.* **2023**, *4*, 80–94.
62. Janzakova, K.; Ghazal, M.; Kumar, A.; Coffinier, Y.; Pecqueur, S.; Alibart, F. Dendritic organic electrochemical transistors grown by electropolymerization for 3D neuromorphic engineering. *Adv. Sci.* **2021**, *8*, 2102973. [CrossRef]
63. Janzakova, K.; Kumar, A.; Ghazal, M.; Susloparova, A.; Coffinier, Y.; Alibart, F.; Pecqueur, S. Analog programing of conducting-polymer dendritic interconnections and control of their morphology. *Nat. Commun.* **2021**, *12*, 6898. [CrossRef]
64. Teka, W.; Marinov, T.M.; Santamaria, F. Neuronal spike timing adaptation described with a fractional leaky integrate-and-fire model. *PLoS Comput. Biol.* **2014**, *10*, e1003526. [CrossRef]
65. Teka, W.W.; Upadhyay, R.K.; Mondal, A. Fractional-order leaky integrate-and-fire model with long-term memory and power law dynamics. *Neural Netw.* **2017**, *93*, 110–125. [CrossRef]