Task-driven Adaptive Sensing on Quadrupole Mass Filter Systems for Classification

Evan X. Chen, Wenzhao Lian, Lawrence Carin, David J. Brady*

Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708

Corresponding author: dbrady@duke.edu

Abstract: An information-theoretical adaptive sensing and classification framework is proposed for Quadrupole mass filter systems. The proposed algorithm designs the most discriminative measurement adaptively by maximizing the mutual information between the class label and the next measurement conditioned on all previous measurements. The proposed adaptive sensing algorithm significantly reduces the number of measurements needed and improves classification accuracy compared with random measurement design. Simulation result on a 76-class mass spectra data library shows a 100% positive detection rate using only 7% adaptive measurements. The reduction of measurements shortens the mass analysis time and theoretically can reduce the required amount of compound material present in the sample for analysis, which potentially increases the sensitivity of the quadrupole mass filter systems.

1. Introduction

Mass spectrometers are widely considered as the most versatile chemical sensors because of their sensitivity and ability to detect a wide range of chemical and biochemical species [1]. Over the past century, mass spectrometer systems have been improved tremendously in the aspect of more diverse ion sources accommodating for different types of materials, various mass analyzers providing different mass range and resolution, and more efficient ion detectors [1, 2]. Most of the above improvements were led by the physics and hardware design. In the last two decades, with an increasing computation power, a few mass spectrometer improvements have been demonstrated by utilizing a combination of hardware modification and computation processing, including increasing system duty cycle of time of flight systems [3], reducing mass analysis time of quadrupole ion trap systems [4], and improving throughput and signal to noise (SNR) of sector systems [5].

Most traditional approaches to mass spectrometer system design involve two separate stages: first, designing the mass spectrometer systems to maximize the fidelity of mass resolution, mass range, SNR, and throughput, and second, designing an algorithm to extract the relevant information for chemical identification/classification from the detected mass spectra. The conventional approaches have shown a disconnection between the spectrometer system design that achieves the maximum mass spectra fidelity and its final task of chemical classification using the collected spectra. In this paper, we propose a computational adaptive sensing framework to address this disconnection through a joint optimization of mass spectrometer design and computer signal processing. The resulting system is designed to achieve the optimal classification performance, rather than the fidelity of mass spectrometer system.

It has been widely known that under certain sampling constraints, compressive sensing can recover certain signals from many fewer measurements than Nyquist rate measurements [6–8]. Furthermore, compared to signal recovery, classification problems requires fewer measurements. A few works exist to design the optimal sensing strategy for classification problems using information-theoretic approaches, including both

non-adaptive [9] and adaptive [10] schemes. Studies have shown that improved performance can be achieved when measurements are adapted the underlying signal of interest [10, 11]. Particularly in [10], an adaptive sensing approach for classification was proposed, *i.e.*, the sensing matrix is designed in a row-wise manner, where each row corresponds to a measurement of a linear projection of the original signal. However, in physical measurement systems, *e.g.*, the quadrupole mass filter system we are considering, the sensing matrix has practical physical constraints, *i.e.*, scanning one point at a time. In this paper, we show an optimal adaptive sensing framework for sequential point scanning classification systems; particularly, we focus on applying this framework to quadrupole mass filter systems. Similar adaptive sensing algorithms for classification have been demonstrated in optical spectroscopy [12] and optimal imaging [13]. These methods have shown significant performance gains compare to conventional systems. However, the adaptive sensing algorithms in [12, 13] are ad-hoc in nature. Specifically, they first designed an optimal sensing matrix without considering the physical system constraints, and then hard-thresholded the optimized sensing matrix to binary afterwards to match the physical system constraints, which leads to suboptimal performances. In contrast, we integrate the physical constraints in the sensing matrix design step, thus guaranteeing the designed adaptive sensing matrix to be optimal.

Our developed sensing algorithm designs discriminative measurements by maximizing the mutual information between the class label and the next measurement, and it continues using the information that has been gathered to adaptively choose the next most informative measurement until reaching a classification confidence level (*e.g.*, 95%). We have applied the proposed algorithm on a 76-class mass spectra library, and our simulation result shows a 100% positive detection rate using only 7% adaptive measurements. Compared with random measurement design, the adaptive algorithm demonstrates significant reduction in the number of measurements needed and classification accuracy improvement. The reduction in the number of required measurements shortens the mass analysis time drastically and theoretically can reduce the required amount of compound material present in the sample for analysis, which potentially increases the sensitivity of the quadrupole mass filter systems.

2. Quadrupole Mass Filter Operation Theory

A quadrupole mass filter system is a type mass analyzer widely used in chemistry and biology for compound classification. It consists of four parallel metal rods as shown in Fig. 1; each opposing pair is connected together, and a radio frequency signal as in (1) is applied to the two pairs of electrode rods,

$$V = U_{DC} + V_{RF}\cos(\Omega t). \tag{1}$$

As a result, the motion of ions in both x and y directions inside the filter follows the solutions of the Mathieu equation,

$$\frac{\partial^2 u}{\partial \tau^2} + (a_u - 2q_u\cos(2\tau))\,u = 0, \tag{2}$$

$$a = a_x = -q_y = \frac{4eU_{dc}}{mr_0^2\Omega^2}, \qquad q = q_x = -q_y = \frac{-2eV_{rf}}{mr_0^2\Omega^2}, \tag{3}$$

where r_0 is the distance from the central axis to the surface of any electrode. The ion's motion in z direction is unaffected by the potentials on the electrodes. The solution of Mathieu equation has two forms: (*i*) periodic and unstable, and (*ii*) periodic and stable. Physically, a stable solution corresponds to a case where the displacement of the particle along either x or y direction is finite, and the associated ion particles entering the mass filter would be transmitted through the device and finally detected by the detector. In contrast, the ion particles corresponding to the unstable solutions have an unstable trajectory and eventually are filtered out of the ion beam due to collision with electrodes. The different m/z ions can be divided into stable and

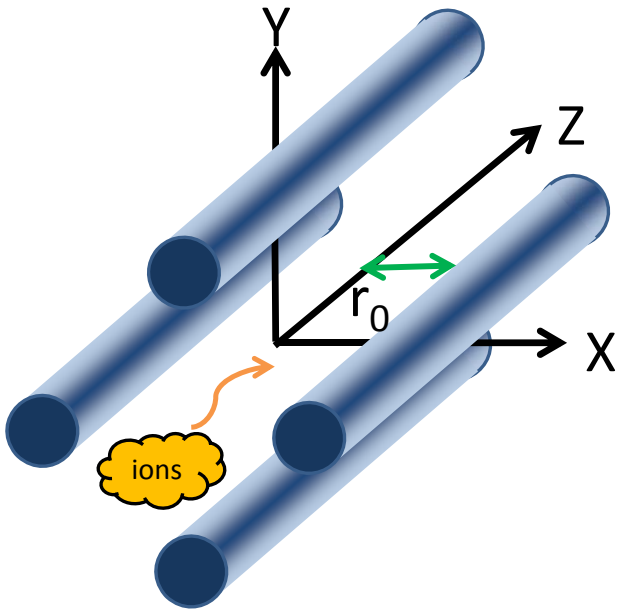

Fig. 1: Schematic of quadrupole mass filter systems.

unstable regions based on a and q values, illustrated by a stability diagram in Fig. 2. The shaded area of the stability diagram is the region where ion particles are stable in both x and y directions.

The quadrupole mass filter operates in the very top apex region of the shaded area of the stability diagram, where $a_{apex} = 0.237$ and $q_{apex} = 0.706$. The applied signal V is designed such that only a small band of m/z is stable and gets through the filter, thus detected by the ion detector. We then rearrange (3) as

$$m/e = \frac{4U_{dc}}{a_{apex} r_0^2 \Omega^2}, \qquad m/e = \frac{-2V_{rf}}{q_{apex} r_0^2 \Omega^2}. \tag{4}$$

To design one m/z which falls in $\{a_{apex}, q_{apex}\}$, we need to fix the ratio of applied U_{dc} and V_{rf} as

$$\frac{U_{dc}}{V_{rf}} = \frac{-a_{apex}}{2q_{apex}}. \tag{5}$$

Without changing the ratio of U_{dc} and V_{rf} as in (5), the amplitude of U_{dc} and V_{rf} are sweeped sequentially to scan the full mass spectrum. The ion signal registered at the detector as function of time is given as

$$y(t) = x[m/e(t)] = x\left[\frac{4U_{dc}(t)}{a_{apex} r_0^2 \Omega^2}\right], \;\; s.t. \;\; \frac{U_{dc}(t)}{V_{rf}(t)} = \frac{-a_{apex}}{2q_{apex}}, \tag{6}$$

where x denotes the abundance intensity. We only presented a abbreviated quadrapole mass filter operation theory, and we refer to [14] and [15] for more details on the topic.

As described, conventional quadrupole mass filter analyzers scan a range of applied voltage sequentially to produce a full mass spectrum, and perform classification in a separate stage. However, the quadrupole mass filter can operate in a more flexible fashion, where a set of discrete signals can be applied to obtain a sequence of discrete m/z measurements. The operation flexibility of quadrupole mass filters provides us an opportunity to apply more smartly designed sensing algorithms on them, *e.g.*, adaptive sensing.

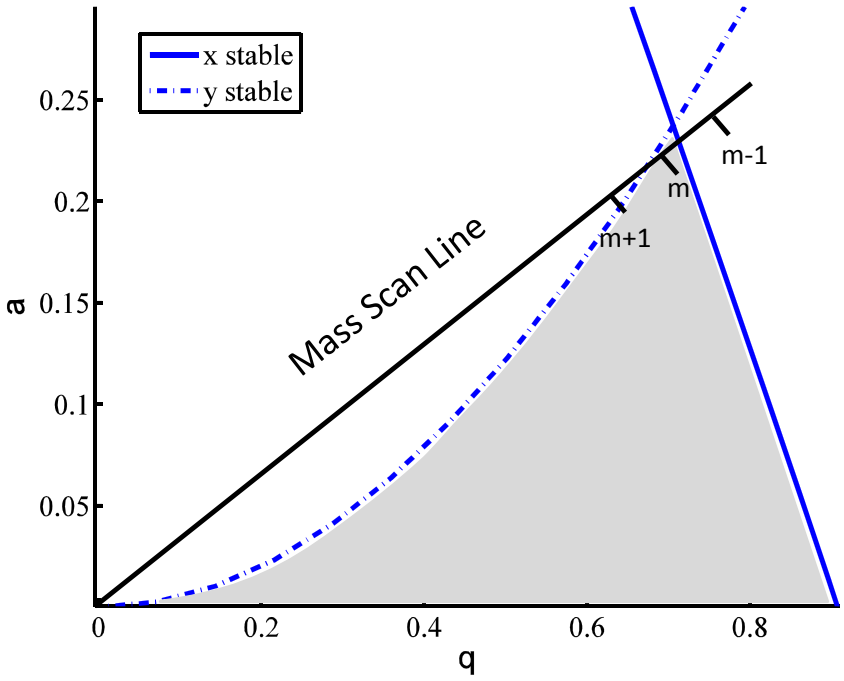

Fig. 2: Stability diagram of quadrupole mass filter systems.

3. Adaptive Sensing Model

3.1. Sensing Matrix Design via Maximizing Mutual Information

We are interested in classifying G classes of compounds, and have a set of training data for each compound class. We assume the class labels and features are generated *i.i.d.* via the following process: $g \sim Mult(g,1,\boldsymbol{w})$, where $\boldsymbol{w} \in \mathbb{R}^{G\times 1}$ is the prior mean on the G classes, and $\boldsymbol{x}|g \sim p(\boldsymbol{x}|g)$, where $\boldsymbol{x} \in \mathbb{R}^{p\times 1}$ is the original signal feature. $p(\boldsymbol{x}|g)$ is the signal distribution for class g, and the marginal distribution for an observed signal is

$$p(\boldsymbol{x}) = \sum_{g=1}^{G} p(g)p(\boldsymbol{x}|g). \tag{7}$$

We then model the class-dependent distributions $p(\boldsymbol{x}|g)$ as multivariate Gaussian

$$p(\boldsymbol{x}|g) = N(\boldsymbol{x}|\boldsymbol{\mu}_g,\boldsymbol{\Sigma}_g) = \frac{1}{\sqrt{(2\pi)^p|\boldsymbol{\Sigma}_g|}}\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_g)^T\boldsymbol{\Sigma}_g^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_g)\right), g \in \{1,2,...,G\}, \tag{8}$$

where $\boldsymbol{\mu}_g \in \mathbb{R}^{p\times 1}$ is the mean and $\boldsymbol{\Sigma}_g \in R^{p\times p}$ is the covariance matrix of the g-th class Gaussian distribution. The discretized version of (6) can be formalized as,

$$\boldsymbol{y} = \boldsymbol{\Phi x} + \boldsymbol{n} \tag{9}$$

As in (9), we assume that $\boldsymbol{n}$ is zero-mean white Gaussian noise, $\boldsymbol{n} \sim N(\boldsymbol{0},\sigma^2\boldsymbol{I})$. Thus, the measurement follows a distribution as $p(\boldsymbol{y}|\boldsymbol{x}) = N(\boldsymbol{y};\boldsymbol{\Phi x},\sigma^2\boldsymbol{I})$. In task-driven classification problems, we seek a sensing

matrix $\mathbf{\Phi}$ such that the sensed signal $\boldsymbol{y}$ is most informative/discrimitive identifying the underlying class label g.

In compressive sensing, we desire to design $\mathbf{\Phi} \in \mathbb{R}^{d\times p}$, where $d \ll p$, to minimize the costs (*e.g.*, time, energy, and money) associated with each measurement. We utilize the mutual information as the feature selection criteria in the sensing matrix design. Mutual information (10) is a measure of how dependent of two random variables are, which can be viewed as the reduction of the uncertainty about $\boldsymbol{x}$ by revealing $\boldsymbol{y}$,

$$I(\boldsymbol{x};\boldsymbol{y}) = -\int\int p(\boldsymbol{x},\boldsymbol{y}) ln\left(\frac{p(\boldsymbol{x})p(\boldsymbol{y})}{p(\boldsymbol{x},\boldsymbol{y})}\right) d\boldsymbol{x}d\boldsymbol{y}. \tag{10}$$

In the signal sensing scenario, where singal recovery is desired, the optimal sensing matrix $\mathbf{\Phi}$ can be designed by finding $argmax_{\mathbf{\Phi}} I(\boldsymbol{x};\boldsymbol{y})$. In our case, we are concerned about signal classification, and the optimal sensing matrix $\mathbf{\Phi}$ can be obtained by finding

$$\mathbf{\Phi} = argmax_{\mathbf{\Phi}} I(g;\boldsymbol{y}), \qquad s.t. \quad \mathbf{\Phi} \text{ matches physical constraint.} \tag{11}$$

Designing the optimal $\mathbf{\Phi}$ with physical constraints is important; otherwise, $\mathbf{\Phi}$ is theoretically optimal, but it's difficult/impossible in actual execution in real physical measurement systems, thus leading to sub-optimal performances. In this paper, we are concerned about sequential point scanning measurement systems, *e.g.*, the quadrupole mass filters. Denote $\mathbf{\Phi} = [\boldsymbol{\phi}_1^T \ ... \boldsymbol{\phi}_k^T \ ... \ \boldsymbol{\phi}_d^T]^T$, where $\boldsymbol{\phi}_k$ is the k-th row of the sensing matrix. The physical constraint of sequential point scanning measurement systems is $\boldsymbol{\phi}_k \in \{0,1\}^{1\times p}$ and $\sum_{i=1}^{p} \phi_{ik} = 1$ for $k \in \{1,2,...,d\}$. Specializing for the sequential point measurement systems, we develop this information-theoretical adaptive sensing algorithm for classification.

Let $\boldsymbol{y}_{(k-1)} = \mathbf{\Phi}_{(k-1)}\boldsymbol{x}$ be the previous (known) $k-1$ measurements, where $\mathbf{\Phi}_{(k-1)} = [\boldsymbol{\phi}_1^T \ ... \boldsymbol{\phi}_{k-1}^T]^T$, and $y_k = \boldsymbol{\phi}_k \boldsymbol{x}$ be the (unknown) next measurement. We want to find the next most informative $\boldsymbol{\phi}_k$ that maximizes the mutual information between the new measurement and the unknown class $g \in \{1,...,G\}$, conditioned on all previous measurements, as

$$\boldsymbol{\phi}_k = argmax_{\boldsymbol{\phi}_k} I(g; y_k | \boldsymbol{y}_{(k-1)}). \tag{12}$$

The mutual information in (12) can be expanded as [10]

$$I(y_k; g|\boldsymbol{y}_{(k-1)}) = H(p(\boldsymbol{y}_{(k)})) - H(p(\boldsymbol{y}_{(k)}|g)) - [H(p(\boldsymbol{y}_{(k-1)})) - H(p(\boldsymbol{y}_{(k-1)}|g))], \tag{13}$$

where $H(\cdot)$ denotes the entropy. The last two terms in (13), $H(p(\boldsymbol{y}_{(k-1)}))$ and $H(p(\boldsymbol{y}_{(k-1)}|g))$, do not depend on $\boldsymbol{\phi}_k$, thus considered as constants and ignored in the optimization step. The first two terms, $H(p(\boldsymbol{y}_{(k)})) - H(p(\boldsymbol{y}_{(k)}|g))$, need to be maximized. The entropy $H(p(\boldsymbol{y}_{(k)}|g)$ has a closed-form expression [10, 16],

$$H(p(\boldsymbol{y}_{(k)}|g) = \frac{1}{2}\left(k(1+\log(2\pi)) + \log|\mathbf{\Sigma}_{\boldsymbol{y}_{(k)}|g}|\right), \tag{14}$$

where $\mathbf{\Sigma}_{\boldsymbol{y}_{(k)}|g} = \mathbf{\Phi}_{(k)}\mathbf{\Sigma}_g\mathbf{\Phi}_{(k)}^T + \sigma^2\boldsymbol{I}$. But the entropy $H(p(\boldsymbol{y}_{(k)}))$ does not yield a closed form expression. We instead use its upper bound as a surrogate [10, 17],

$$H(p(\boldsymbol{y}_{(k)})) \leq \frac{1}{2}\left(k(1+\log(2\pi)) + \log|\bar{\mathbf{\Sigma}}_{\boldsymbol{y}_{(k)}}|\right), \tag{15}$$

where $\bar{\mathbf{\Sigma}}_{\boldsymbol{y}_{(k)}} = \mathbf{\Phi}_{(k)}\bar{\mathbf{\Sigma}}\mathbf{\Phi}_{(k)}^T + \sigma^2\boldsymbol{I}$, $\bar{\mathbf{\Sigma}} = \sum_{g=1}^{G} p(g)(\mathbf{\Sigma}_g + (\bar{\boldsymbol{\mu}} - \boldsymbol{\mu}_g)(\bar{\boldsymbol{\mu}} - \boldsymbol{\mu}_g)^T)$, and $\bar{\boldsymbol{\mu}} = \sum_{g=1}^{G} p(g)\boldsymbol{\mu}_g$.

Thus, the optimal k-th row of the sensing matrix $\boldsymbol{\phi}_k$ can be obtained by solving (13). Because there is a limited number of choices on $\boldsymbol{\phi}_k$ in our point scanning measurement systems, at the k-th step, we can plug in all possible $\boldsymbol{\phi}_k$ into equation (13) and choose the one yielding the maximum conditional mutual information

between class label g and y_k given $\boldsymbol{y}_{(k-1)}$. Then the current belief of label distribution $p(g)$, *i.e.*, $p(g|\boldsymbol{y}_{(k-1)})$, is updated to $p(g|\boldsymbol{y}_{(k)})$, *i.e.*, the label distribution after including the newly acquired measurement y_k using the Bayes' theorem (16).

$$p(g|\boldsymbol{y}_{(k)}) = \frac{p(y_k|g)p(g|\boldsymbol{y}_{(k-1)})}{\sum_{i=1}^{G} p(y_k|g=i)p(g=i|\boldsymbol{y}_{(k-1)})}. \tag{16}$$

Then in the next step ($k = k+1$), the most informative measurement can be designed by solving (12) again based on the updated model. The above adaptive sensing design and classification procedure continuous until *max* $p(g|\boldsymbol{y}_{(k)})$ reaches a confidential threshold (here used 95%) and finally the class label g which yield the *max* $p(g|\boldsymbol{y}_{(k)})$ is assigned to the unknown measurement.

3.2. Multivariate Gaussian Distribution Estimation by Probabilistic Principal Component Analysis

Given a set of features $\boldsymbol{X} = (\boldsymbol{x}_1, ..., \boldsymbol{x}_N)^T$ belonging to the same class g, the observations $\{\boldsymbol{x}_n\}$ are assumed to be drawn independently from a multivarate Gaussian distribution, and we can estimate the parameters of (8), *i.e.*, $\boldsymbol{\mu}_g$ and $\boldsymbol{\Sigma}_g$, via the maximum likelihood approach. The maximum likelihood estimation (MLE) of mean and covariance are computed as [18],

$$\boldsymbol{\mu}_{MLE} = \frac{1}{N}\sum_{n=1}^{N} \boldsymbol{x}_n, \quad \boldsymbol{\Sigma}_{MLE} = \frac{1}{N}\sum_{n=1}^{N} (\boldsymbol{x}_n - \boldsymbol{\mu}_{ML})(\boldsymbol{x}_n - \boldsymbol{\mu}_{ML})^T. \tag{17}$$

The MLE of mean and covariance matrix of multivariate Gaussian distribution works in sufficient data scenarios, *i.e.*, $N \gg p$. However, in the large p small N setting or the covariance matrix has a low rank structure, using this direct approach, the estimated covariance matrix is ill-conditioned.

In the scenario we consider, the covariance matrix has a low rank structure, *i.e.*, only a small number of principle components have significant variance. Therefore, to mitigate the distribution estimation problem. we use the probabilistic principal component analysis (PPCA) [19]. Specifically, the PPCA constructs the following model for each multivariate Gaussian distribution

$$p(\boldsymbol{x}|\boldsymbol{t}) \sim N(\boldsymbol{W}\boldsymbol{t} + \boldsymbol{\mu}, \sigma^2\boldsymbol{I}), \tag{18}$$

where $\boldsymbol{W} \in \mathbb{R}^{p\times q}$ relates the observed variable x and latent variable $\boldsymbol{t} \in \mathbb{R}^q$, with $d < p$. The latent variable $\boldsymbol{t}$ offers a more parsimonious explanation of the dependencies between observations. The distrubtion over the latent variables is also Gaussian defined as $\boldsymbol{t} \sim N(\boldsymbol{0}, \boldsymbol{I})$. Then the marginal distribution for the observed data $\boldsymbol{x}$ can be obtained by integrating out the latent variables $\boldsymbol{t}$ as,

$$\boldsymbol{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}_{ppca}), \tag{19}$$

where $\boldsymbol{\Sigma}_{ppca} = \boldsymbol{W}^T\boldsymbol{W} + \sigma^2\boldsymbol{I}$. Now (19) serves as the distribution for each class, we need to estimate the distribution parameters including $\boldsymbol{\mu}$, $\boldsymbol{W}$ and σ^2. The MLE of $\boldsymbol{\mu}$ is the same as in (17). For $\boldsymbol{W}$, the MLE can be derived as [19],

$$\boldsymbol{W}_{MLE} = \boldsymbol{U}_q(\boldsymbol{\Lambda}_q - \sigma^2\boldsymbol{I})^{1/2}, \tag{20}$$

where $\boldsymbol{U}_q \in \mathbb{R}^{p\times q}$ are the principal eigenvectors of $\boldsymbol{\Sigma}$, and $\boldsymbol{\Lambda}_q \in \mathbb{R}^{q\times q}$ is a diagonal matrix with its diagonal entries corresponding to eigenvalues $\lambda_1, ..., \lambda_q$ of $\boldsymbol{\Sigma}$. Lastly, the maximum likelihood estimation of σ^2 is given by [19]

$$\sigma^2_{MLE} = \frac{1}{p-q}\sum_{j=q+1}^{p} \lambda_j. \tag{21}$$

In above, $\boldsymbol{W}_{MLE}$ reflects a mapping from the latent variable space into the principal component subspace of the observed data space. σ^2_{MLE} can be interpreted as the residual averaged over the "lost" principal component dimensions.

4. Results and Discussion

We applied our proposed adaptive sensing algorithm on a 76 chemical mass spectrum library, where each signature (signal vector) has a dimension of $p = 373$ m/z channels (features). A mass spectrometer system measures discrete counts of ions, and the measured spectrum can is assumed as Poisson distributed following convention [20]. Due to the limit access to real mass spectrum measurements, we simulated 1000 realizations of each class from a Poisson distribution with the real chemical mass spectrum signature as the mean. We divide the data into 99% for training and the rest 1% for classification testing. The Poisson distributed data can be reasonably modeled by a Gaussian distribution when the Poisson mean is high [21], and we directly apply the proposed algorithm to this dataset. In training, the class-dependent distributions are learned via PPCA.

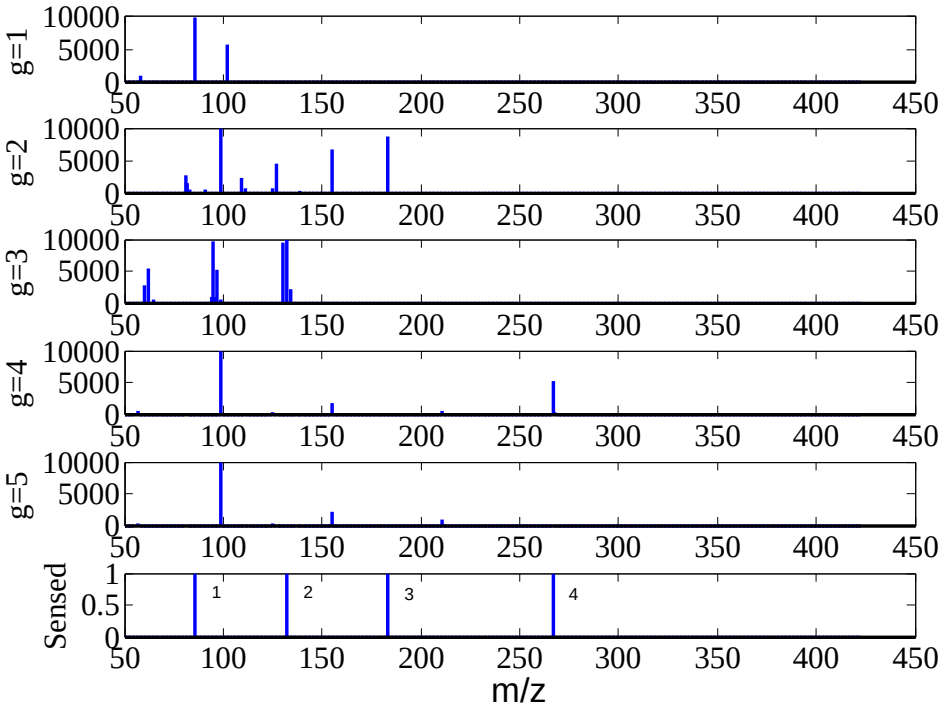

Fig. 3: Spectra signature of 5 chemical classes and the designed adaptive measurements for one realization from class 5.

Fig. 3 illustrates 5 chemical class signatures and the corresponding optimally designed sensing points for a chemical realization from class 5. First, the algorithm obtains a measurement at $m/z = 87$, calculates the posterior of label distribution, and eliminates the possibility of class 1. The algorithm continues the posterior update and informative measurement design until the posterior converges (*e.g.*, $p(g = \hat{g}) > 0.95$ for some $\hat{g}$). As seen from Fig. 3, the second and third measurements eliminate the possibility of class 3 and 2, respectively. After the fourth adaptively designed measurement, the algorithm classifies the unknown chemical as in class 5. As seen from this simple example, the proposed algorithm can identify all 5 classes by using only a maximum of 4 measurements. In general, the needed number of measurements grows as the

number of classes increases.

The result of applying the proposed algorithm on the 76-class full dataset is shown in Fig. 4, where the percentage of positive detection (*Pd*) is plotted as a function of the number of measurements. Mean and standard deviation over five random training and testing data partitions are reported. We observe that *Pd* increases monotoniclly as more measurements are obtained, until it achieves a 100% *Pd* with 26 measurements, which is about 7% of the total channels measured. Compared with the randomly selected measurements from all m/z channels, our proposed adaptive sensing algorithm performs significantly better, both in terms of accuracy and the number of measurements needed. The empirical distribution of the number of measurements needed for this 76-class classification task is shown in Fig. 5. The adaptive sensing algorithm used a maximum at 26 and mean at 12 measurements to achieve 100% Pd. The histogram is left skewed, which indicates a large portion of testing samples require fewer measurements than the average for successful classification.

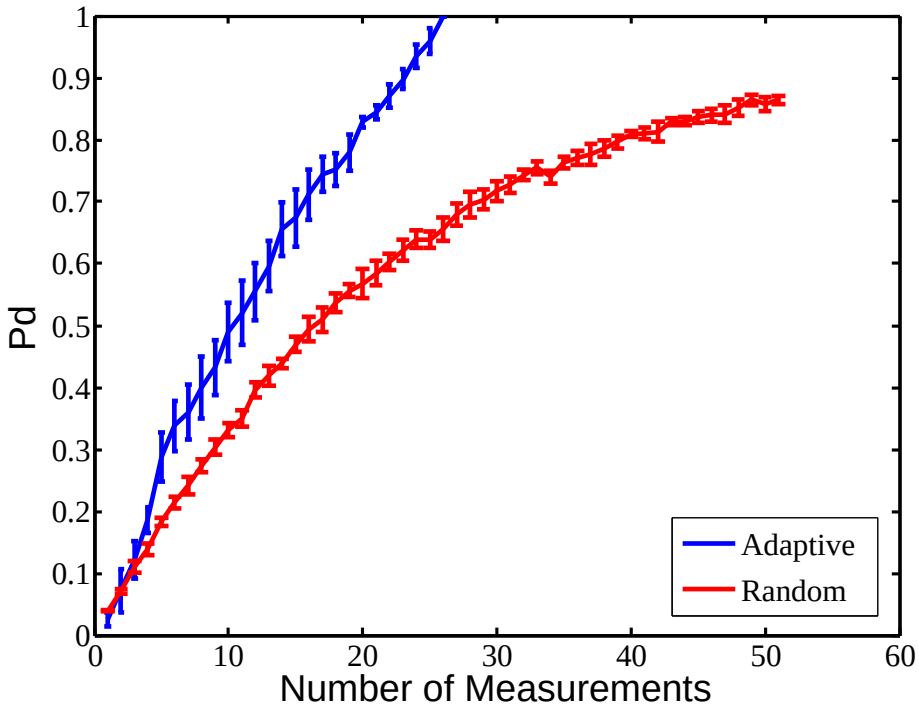

Fig. 4: 76-class classification results.

Quadrupole mass filter is coupled with a continuous ion source. It requires a continues supply of compound material and ionization until the systems finishes the mass analysis. Our proposed adaptive sensing algorithm achieves 100% *Pd* using 7% adaptive measurements, which reduces the mass analysis time by a factor of 14. This can lead to a significant reduction on the amount of required material in the sample, thus yielding system sensitivity improvement. Alternatively, due to the reduction of the number of measurements, we can allocate more time on each informative mass spectrum channel measurement to achieve better SNR on the mass spectrometer system.

In the above experiment, we have demonstrated the application of the adaptive sensing algorithm on quadrupole mass filter system. However, the algorithm can be applied to any other sequential point scanning measurement systems for the purpose of classification, including high-field asymmetric waveform ion mobility spectrometers (FAIMS) [22] and differential mobility spectrometers (DMS) [23]. Similar to

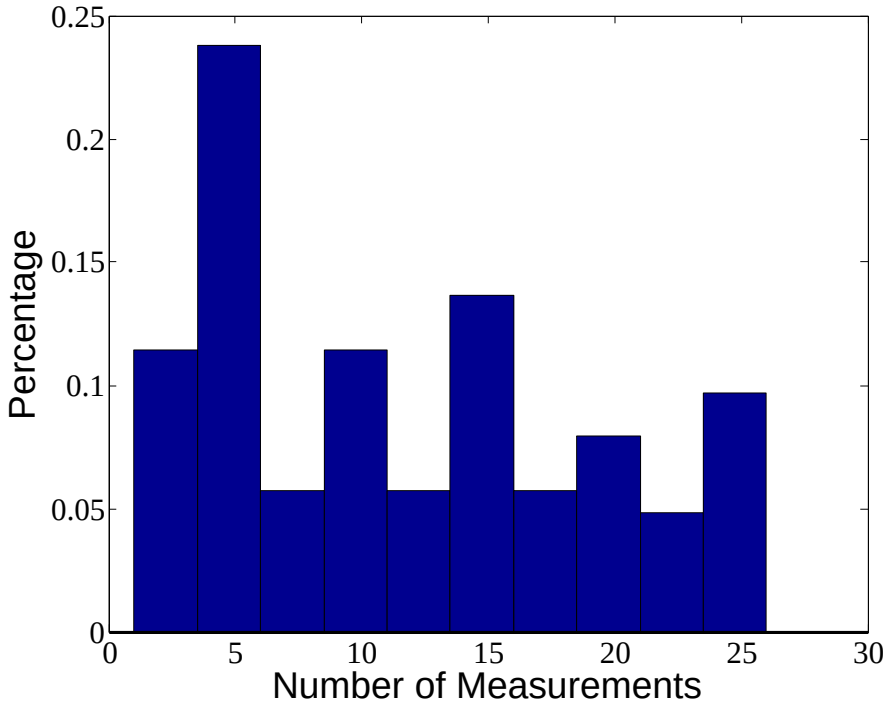

Fig. 5: Number of measurements histogram.

quadrapole mass filter systems, both FAIMS and DMS operate as bandpass filters, and a RF voltage is sweeped sequentially to measure the full spectrum [24] and then followed with classification analysis. Therefore, the proposed adaptive sensing algorithm can be readily applied to both FAIMS and DMS systems to reduce analysis time and potentially improve sensitivity of the spectrometers.

5. Conclusion

We developed a task-driven adaptive sensing framework on sequential point scanning measurement system for classification, and demonstrated its application on quadrupole mass filter analyzer systems for faster mass analysis. Our simulation results have demonstrated a 100% *Pd* positive detection rate by using only a maximum of 7% adaptive measurements. The significant measurement reduction can lead to system throughput or SNR improvements. Future work includes performing adaptive quadruple mass filter experiments in physical system and exploring the applications of the proposed algorithm to asymmetric waveform ion mobility spectrometer and differential mobility spectrometers.

References

1. J. H. Gross, *Mass spectrometry: a textbook* (Springer Science & Business Media, 2004).
2. S. A. McLuckey and J. M. Wells, "Mass analysis at the advent of the 21st century," Chemical reviews **101**, 571–606 (2001).
3. A. Brock, N. Rodriguez, and R. Zare, "Hadamard transform time-of-flight mass spectrometry," Analytical Chemistry **70**, 3735–3741 (1998).

4. E. Chen, M. Gehm, R. Danell, M. Wells, J. Glass, and D. Brady, "Compressive mass analysis on quadrupole ion trap systems," Journal of The American Society for Mass Spectrometry **25**, 1295–1304 (2014).
5. Z. Russell, E. Chen, J. Amsden, S. Wolter, R. Danell, C. Parker, B. Stoner, M. Gehm, D. Brady, and J. Glass, "Two-dimensional aperture coding for magnetic sector mass spectrometry," Journal of The American Society for Mass Spectrometry **26**, 248–256 (2015).
6. D. J. Brady, *Optical imaging and spectroscopy* (Wiley-OSA, 2009).
7. E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," Information Theory, IEEE Transactions on **52**, 489–509 (2006).
8. D. L. Donoho, "Compressed sensing," Information Theory, IEEE Transactions on **52**, 1289–1306 (2006).
9. M. Chen, W. Carson, M. Rodrigues, R. Calderbank, and L. Carin, "Communications inspired linear discriminant analysis," arXiv preprint arXiv:1206.6397 (2012).
10. J. M. Duarte-Carvajalino, G. Yu, L. Carin, and G. Sapiro, "Task-driven adaptive statistical compressive sensing of gaussian mixture models," Signal Processing, IEEE Transactions on **61**, 585–600 (2013).
11. W. R. Carson, M. Chen, M. R. Rodrigues, R. Calderbank, and L. Carin, "Communications-inspired projection design with application to compressive sensing," SIAM Journal on Imaging Sciences **5**, 1185–1212 (2012).
12. D. Dinakarababu, D. Golish, and M. Gehm, "Adaptive feature specific spectroscopy for rapid chemical identification," Optics Express **19**, 4595–4610 (2011).
13. A. Ashok, P. K. Baheti, and M. A. Neifeld, "Compressive imaging system design using task-specific information," Applied optics **47**, 4457–4471 (2008).
14. P. E. Miller and M. B. Denton, "The quadrupole mass filter: Basic operating concepts," Journal of Chemical Education **63**, 617 (1986).
15. D. L. Shinholt, S. N. Anthony, A. W. Alexander, B. E. Draper, and M. F. Jarrold, "A frequency and amplitude scanned quadrupole mass filter for the analysis of high m/z ions," Review of Scientific Instruments **85**, 113109 (2014).
16. T. M. Cover and J. A. Thomas, *Elements of information theory* (John Wiley & Sons, 2012).
17. M. Padmanabhan and S. Dharanipragada, "Maximizing information content in feature extraction," Speech and Audio Processing, IEEE Transactions on **13**, 512–519 (2005).
18. C. M. Bishop *et al.*, *Pattern recognition and machine learning*, vol. 4 (springer New York, 2006).
19. M. E. Tipping and C. M. Bishop, "Probabilistic principal component analysis," Journal of the Royal Statistical Society: Series B (Statistical Methodology) **61**, 611–622 (1999).
20. P. D. Piehowski, A. M. Davey, M. E. Kurczy, E. D. Sheets, N. Winograd, A. G. Ewing, and M. L. Heien, "Time-of-flight secondary ion mass spectrometry imaging of subcellular lipid heterogeneity: Poisson counting and spatial resolution," Analytical chemistry **81**, 5593–5602 (2009).
21. T. A. Ryan, H. Reuter, and S. J. Smith, "Optical detection of a quantal presynaptic membrane turnover," Nature **388**, 478–482 (1997).
22. R. Guevremont, "High-field asymmetric waveform ion mobility spectrometry: a new tool for mass spectrometry," Journal of Chromatography A **1058**, 3–19 (2004).
23. G. Eiceman, E. Krylov, N. Krylova, E. Nazarov, and R. Miller, "Separation of ions from explosives in differential mobility spectrometry by vapor-modified drift gas," Analytical chemistry **76**, 4937–4944 (2004).
24. B. M. Kolakowski and Z. Mester, "Review of applications of high-field asymmetric waveform ion mobility spectrometry (faims) and differential mobility spectrometry (dms)," Analyst **132**, 842–864 (2007).