ACCELERATED TIME-OF-FLIGHT MASS SPECTROMETRY

Morteza Ibrahimi, Andrea Montanari

Stanford University
Stanford, CA

George S. Moore

Agilent Technologies
Santa Clara, CA

ABSTRACT

We employ a simple modification to the conventional time of flight mass spectrometry (TOFMS) where a *variable* and (pseudo)-*random* pulsing rate is used which allows for traces from different pulses to overlap. This modification requires little alteration to the currently employed hardware. However, it requires a reconstruction method to recover the spectrum from highly aliased traces. We propose and demonstrate an efficient algorithm that can process massive TOFMS data using computational resources that can be considered modest with today's standards. Our approach can be used to improve duty cycle, throughput, and mass resolution of TOFMS at the same time. We expect this to extend the applicability of TOFMS to new domains.

1. INTRODUCTION

Mass spectrometry (MS) refers to a family of techniques used to analyze the constituent chemical species in a sample. The applications abound in science and technology and new fields of scientific investigations have evolved around these techniques. Time of flight (TOF) mass spectrometry was introduced in the 1940s by Stephens [1]. A basic TOF mass spectrometer consists of three parts: a source region, a drift region, and a detector. In the source region, the input is ionized and subsequently accelerated by a static electrical field into the drift region. Ideally, the ions entering the drift region have kinetic energy proportional to their charge. Hence, the time that takes for an ion to reach the detector is proportional to $\sqrt{m/z}$ where m is the mass of the ion and z is its net charge.

TOFMS offers two major benefits over alternative techniques. It has essentially unlimited mass range and high repetition rate. These properties along with the recent advances in available hardware and ionization techniques have made TOFMS an appealing choice for the analysis of samples with wide mass range [2], biological macromolecules [3] and in combination with other mass spectrometers [4]. However, there exists an intrinsic trade-off between throughput (number of samples that can be analyzed in a given time) and duty cycle (fraction of ions accelerated into the drift region from a continuous ion source) on one hand, and mass resolution (the smallest difference in m/z that can be resolved) and mass range on the other hand. Furthermore, some applications have stringent requirements in terms of throughput and duty cycle, e.g., when spectrometers are used in tandem, or preceded by chromatographic separation. Therefore, simultaneous improvement of the throughput, duty cycle, and mass resolution is of fundamental interest in TOFMS [5].

TOFMS is a pulsed technique, i.e., ions are formed in an ionization stage and subsequently accelerated as a *packet* into the drift region. The ion species in each packet travel through the drift region with different speeds and arrive at the detector at different times. As the ions impact the detector they generate a continuous electrical signal which is then sampled resulting in a discrete signal. The result of this process, which we call an *acquisition*, is a noisy sample of the $\sqrt{m/z}$ spectrum. A single acquisition is often too noisy and this process is repeated from hundreds to a few thousands times and averaged to obtain an accurate estimate of the spectrum. In conventional TOFMS, the time between consecutive pulses is set to be long enough to avoid overlap between different acquisitions, i.e., for the slowest ion in an acquisition to arrive at the detector before the fastest ion of the next acquisition. This fact ties the mass resolving power of TOFMS to the duty cycle and throughput; a longer drift region results in higher mass resolution but decreases the throughput and duty cycle.

There has been previous work trying to alleviate this problem. One approach, called Fourier transform TOF, is to modulate a continuous ion beam at the source using a periodic waveform and subsequently accelerate it into the drift region [6]. The detected signal is then demodulated to obtain the spectrum. Another approach, called Hadamard transform TOF (HT-TOF) [7], is based on modulation (gating) of a continuous ion source by a $0/1$ pulse. The spectrum is obtained by a deconvolution that can be implemented efficiently using the fast Hadamard transform.

One drawback of these methods is that they treat the reconstruction process as a deterministic inversion problem and ignore the noisy nature of the observations. Furthermore, they require substantial modification to the hardware of a conventional TOFMS. In this paper, we describe a different method, called Accelerated TOF (A-TOFMS), which simultaneously achieve mass resolution, duty cycle, and throughput improvement using essentially the same hardware as a conventional TOFMS. Our reconstruction scheme acknowledges the stochastic nature of the observation. Simulation results using real data confirm the performance improvement of this scheme.

Notations and Terminology: Let $\boldsymbol{x}[t]$, $t = 1, 2, \dots, n$ be the output of the detector for a single acquisition. We refer to $\boldsymbol{x}[t]$ as to a *transient*. Typically a TOFMS experiment consists of many acquisitions which are later processed (simply averaged) to obtain a more accurate estimate of the spectrum. Let $\boldsymbol{x}^{(l)}[t]$ be the l^{th} transient. Define the *true spectrum*, $\bar{\boldsymbol{x}}[t]$, as the average of infinitely many transients, i.e., $\bar{\boldsymbol{x}}[t] = \lim_{L\to\infty} \frac{1}{L}\sum_{l=1}^{L} \boldsymbol{x}^{(l)}[t]$. Each transient $\boldsymbol{x}^{(l)}[t]$ is therefore a noisy version of $\bar{\boldsymbol{x}}[t]$. We define the *trace*, $\boldsymbol{y}[t]$, to be the observed detector response for multiple, possibly overlapping acquisition. Given an observed trace $\boldsymbol{y}[t]$, the goal is to find a *good* estimate $\widehat{x}[t]$ of $\bar{\boldsymbol{x}}[t]$. In what follows we treat the discrete signals as column vectors, and we let v^* denote the transpose of v and $\langle u, v\rangle$ denote the dot product of u and v.

When an ion impacts the detector it generates a bell-shaped pulse in the output of the detector. We refer to an observed pulse in the trace as an *impact event*, or event for short. Usually the sampling rate of the detector response is such that an event spans multiple samples. For pedagogical reasons, we will describe the algorithm *as*

if each event could occupy only one sample and there was no time jitter, i.e., all the ions of the same species were observed in the same bin. The algorithm can be applied without this assumptions with some minor modifications and all the results presented here are for the general case.

2. MEASUREMENT SCHEME AND THE RECONSTRUCTION ALGORITHM

A TOF measurement from a single acquisition is commonly very sparse (after removing the additive electrical noise through preprocessing, c.f. Section 4). Furthermore, a single measurement of the whole spectrum is not expensive and it can be viewed as being performed in parallel as all ions are flying in the drift region at the same time. However, the observed signal from a single acquisition is too noisy and many repetitions of the same measurement are necessary to obtain an accurate estimate of the spectrum. In a conventional TOFMS setting the observed trace can be expressed as $\boldsymbol{y}[t] = \sum_{l=1}^{L} \boldsymbol{x}^{(l)}[t - ln]$ where $\boldsymbol{x}^{(l)}$ is the detector response to the l^{th} acquisition and $\boldsymbol{x}[t]$ is understood to be zero for $t \leq 0$ or $t > n$.

We incorporate a simple, yet powerful, modification to this scheme [8]. Define τ_l, the *firing time*, to be the starting time of the l^{th} transient, i.e., the time when the l^{th} ion packet is accelerated into the drift region. Define $\Delta\tau_l \equiv \tau_{l+1} - \tau_l$. In conventional TOFMS $\Delta\tau_l = \Delta\tau \geq n$ to avoid overlapping between consecutive transients. We relax this condition and let $\Delta\tau_l$ be a random variable with $\mathbb{E}[\Delta\tau_l] = \alpha n$, for some $\alpha < 1$. As is the case with the HT-TOF we assume that the detector response to overlapping acquisitions is the superposition of the individual responses,

$$\boldsymbol{y}[t] = \sum_{l=1}^{L} \boldsymbol{x}^{(l)}[t - \tau_l]. \tag{1}$$

In this case at each time t, $\boldsymbol{y}[t]$ is the superposition of multiple overlapping transients. Assume there are a total of L transients and let $0 = \tau_1 < \tau_2 < \cdots < \tau_L$ be the firing times. For a given $\boldsymbol{\tau} = (\tau_1, \tau_2, \ldots, \tau_L)$, define the adjacency matrix $A \in \mathbb{R}^{T \times n}$ as

$$A_{ti} = \begin{cases} 1 & \text{if } \exists\ l \in \{1, 2, \ldots, L\} \text{ s.t. } i = t - \tau_l \\ 0 & \text{Otherwise.} \end{cases} \tag{2}$$

Then, $\boldsymbol{y}[t]$ can be considered as a noisy version of linear measurements of $\bar{\boldsymbol{x}}$, $\langle A_t, \bar{\boldsymbol{x}} \rangle$, with A_t the t^{th} row of A. In this notation, the conventional TOFMS is a special case where each row of A has only one nonzero element, i.e., measurement $\boldsymbol{y}[t]$ corresponds to a noisy observation of $\bar{\boldsymbol{x}}[i]$ for some bin i. The structure of matrix A reveals the difference between A-TOFMS and conventional TOFMS.

$$A_{TOF} = \begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \\ 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \\ 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \end{pmatrix}, \quad A_{ATOF} = \begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 1&0&0&1 \\ 0&1&0&0 \\ 1&0&1&0 \\ 0&1&0&1 \\ 1&0&1&1 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \end{pmatrix}$$

Given the trace $\boldsymbol{y}$ and adjacency matrix A one can attempt to solve for $\bar{\boldsymbol{x}}$ using an ordinary least squares, $\hat{\boldsymbol{x}}_{LS} = \text{argmin}\|A\boldsymbol{x} - y\|_2$ or ℓ_1-regularized least squares $\hat{\boldsymbol{x}}_{LASSO} = \text{argmin}\|A\boldsymbol{x} - y\|_2 + \lambda\|\boldsymbol{x}\|_1$. However, simulation results demonstrate poor performance for both these methods. The reason lies in the choice of the objective function. Sum of square residuals approximates the negative log likelihood when the measurement noise is additive Gaussian. However, TOFMS is dominated by shot-noise which is signal dependent and non-additive. Similar issues arises in applications like photon-limited imaging where the observation are again shot-noise limited. Regularized maximum likelihood approaches proved effective in these settings [9]. Here we propose a stochastic model for the observation $\boldsymbol{y}$ and present an algorithm that optimizes the ℓ_1-regularized log likelihood.

Define $\boldsymbol{w} \in \mathbb{R}^n$ such that $\boldsymbol{w}[i]$ is the average number of ions that impact the detector at bin i for a single acquisition. It is natural to assume that the number of ions that impact the detector at time i follows a Poisson distribution with mean $\boldsymbol{w}[i]$. However, experimental results indicate that a Poisson model for $\boldsymbol{y}$ is inadequate. Hence, we assume each ion impact generates an ADC count which itself is an exponentially distributed random variable. Therefore, conditioning on k ions impacting the detector at a certain time the ADC count has an Erlang distribution with the shape parameter k. This model conforms both with experimental data and current understanding of the underlying mechanism by which ion detectors function.

Let μ be the mean of the exponential random variable describing the detector response. Given $\boldsymbol{w}$ and A the probability density function of $\boldsymbol{y}[t]$ is

$$\mathbb{P}(\boldsymbol{y}[t]|\boldsymbol{w}, A) = e^{-\langle A_t, \boldsymbol{w} \rangle}\delta(\boldsymbol{y}[t]) + \sum_{k=1}^{\infty} E_{k,\mu}(\boldsymbol{y}[t]) P_{\langle A_t, \boldsymbol{w} \rangle}(k),$$

where $E_{k,\mu}(\,\cdot\,)$ and $P_\lambda(\,\cdot\,)$ are the Erlang pdf and Poisson pmf defined as $E_{k,\mu}(y) = \frac{y^{k-1}\exp(-\frac{y}{\mu})}{\mu^k(k-1)!}$, $P_\lambda(k) = \frac{\exp(-\lambda)\lambda^k}{k!}$, and $\delta(\,\cdot\,)$ is the Dirac delta representing a probability mass at zero. Assuming that each ion impact is observed in only one element of the trace, different trace elements are the result of different ion impacts. Hence, given $\boldsymbol{w}$ and A the observed responses in different times can be considered independent, namely $\mathbb{P}(\boldsymbol{y}|\boldsymbol{w}, A) = \prod_{t=1}^{T} \mathbb{P}(\boldsymbol{y}[t]|\boldsymbol{w}, A)$. Hence the negative log-likelihood function takes the form

$$\ell_{\boldsymbol{y},A}(\boldsymbol{w}) = \langle \boldsymbol{1}, A\boldsymbol{w} \rangle - \sum_{t \in F} \log \left\{ \sqrt{\langle A_t, \boldsymbol{w} \rangle}\ \mathrm{I}_1\left(\frac{2\sqrt{\boldsymbol{y}[t]\langle A_t, \boldsymbol{w} \rangle}}{\sqrt{\mu}}\right) \right\} \tag{3}$$

where I is the modified Bessel Function of type 1 and we dropped the term $\frac{1}{2}(\log(\boldsymbol{y}[t]) - \log(\mu))$ which does not depend on $\boldsymbol{w}$. We note that the negative log-likelihood function (3) is strictly convex in $\boldsymbol{w}$ which is remarkable given the existence of the *hidden* variable k.

3. ALGORITHM

Define the regularized negative log-likelihood cost function $\mathcal{C}_{\boldsymbol{y},A,\lambda}(\boldsymbol{w}) = \ell(\boldsymbol{w}|\boldsymbol{y}, A) + \lambda\|\boldsymbol{w}\|_1$. Our algorithm attempts to solve the following optimization problem.

$$\underset{\boldsymbol{w}}{\text{minimize}} \quad \ell_{\boldsymbol{y},A}(\boldsymbol{w}) + \lambda\|\boldsymbol{w}\|_1\,, \qquad \text{s.t.} \quad \boldsymbol{w} \geq 0, \tag{4}$$

where $\|\boldsymbol{w}\|_1 = \sum_{i=1}^{n} |\boldsymbol{w}[i]|$ is the regularization penalty and λ is a parameter. For a doubly differentiable function $f : \mathcal{D} \subset \mathbb{R}^n \to \mathbb{R}$ let ∇f and $\nabla^2 f$ be the gradient and Hessian of f respectively. Assume that there exist a constant $\gamma > 0$ such that $\|\nabla^2 \ell_{\boldsymbol{y},A}(\boldsymbol{w})\|_2 < \gamma^{-1}$ for $\boldsymbol{w}$ in some set of interest. We call the parameter γ the step size as we use it to scale the steps the algorithm takes in each iteration. Let $\boldsymbol{w}^{(k)}$ be our estimate of $\boldsymbol{w}$ at step k. Then we can compute an upper bound for the cost function $\mathcal{C}_{\boldsymbol{y},A,\lambda}(\boldsymbol{w})$ as follows.

$$\begin{aligned} \mathcal{C}_{\boldsymbol{y},A,\lambda}(\boldsymbol{w}) \leq\ & \ell_{\boldsymbol{y},A}(\boldsymbol{w}^{(k)}) + \boldsymbol{w}^* \nabla \ell_{\boldsymbol{y},A}(\boldsymbol{w}^{(k)}) \\ & + \gamma^{-1}\|\boldsymbol{w} - \boldsymbol{w}^{(k)}\|_2^2 + \|\boldsymbol{w}\|_1. \end{aligned} \tag{5}$$

Equation (5) provides an approximation for $\mathcal{C}_{\boldsymbol{y},A,\lambda}(\boldsymbol{w})$ when $\boldsymbol{w}$ is in a small neighborhood of $\boldsymbol{w}^{(k)}$. Minimizing the right-hand-side of Eq. (5) with respect to $\boldsymbol{w}$ as a surrogate for the actual cost function results in

$$\boldsymbol{w}^{(t+1)} = \eta_\theta \left(\boldsymbol{w}^{(k)} - \gamma \nabla \ell_{\boldsymbol{y},A}(\boldsymbol{w}^{(k)}) \right). \tag{6}$$

where $\eta_\theta(\,\cdot\,)$ is the soft thresholding function, $\eta_\theta(x) = \mathrm{sgn}(x)(|x| - \theta)_+$ with $(\,\cdot\,)_+$ being the positive part and $\theta \propto \lambda^{-1}$. From Eq. (3), $\nabla \ell_{\boldsymbol{y},A}(\boldsymbol{w})$ can be calculated as

$$\nabla \ell^*_{\boldsymbol{y},A}(\boldsymbol{w}) = A^* \mathbf{1} - \sum_{t \in F} \left\{ \frac{1}{2\langle A_t, \boldsymbol{w} \rangle} + \frac{\mathrm{I}_0\left(2\sqrt{\frac{\boldsymbol{y}[t]}{\mu}\langle A_t, \boldsymbol{w}\rangle}\right) + \mathrm{I}_2\left(2\sqrt{\frac{\boldsymbol{y}[t]}{\mu}\langle A_t, \boldsymbol{w}\rangle}\right)}{2\sqrt{\frac{\boldsymbol{y}[t]}{\mu}\langle A_t, \boldsymbol{w}\rangle}\,\mathrm{I}_1\left(2\sqrt{\frac{\boldsymbol{y}[t]}{\mu}\langle A_t, \boldsymbol{w}\rangle}\right)} \right\} A_t. \tag{7}$$

Given the gradient of the log-likelihood the algorithm is as follows.

Algorithm

Input: trace $\boldsymbol{y}$, firing times $\boldsymbol{\tau}$, and constants $(\theta_0, \theta_1, \mu)$
Output: estimated spectrum $\hat{\boldsymbol{w}}$

1: Calculate the adjacency matrix A as in Eq. (2).
2: Set $\boldsymbol{w}^{(0)} = 0$, $\theta = \theta_0 + \theta_1$.
2: Repeat until stopping criterion is met
$\theta \leftarrow \theta_0 + \frac{1}{k^2}\theta_1$
$\boldsymbol{w}^{(k+1)} \leftarrow \eta_\theta \left(\boldsymbol{w}^{(k)} - \gamma \nabla \ell_{\boldsymbol{y},A}(\boldsymbol{w}^{(k)}) \right)$.
3: Return $\hat{\boldsymbol{x}} \equiv \mu \boldsymbol{w}^{(k)}$.

4. EXPERIMENTAL EVALUATION

In this section we present performance evaluation results for the A-TOFMS algorithm. We obtain $10,000$ transients for a high concentration multimode chemical sample using conventional TOFMS technique. Each transient is 100 µs in length sampled at 25 ps intervals. Therefore, in our notation $n = 4 \times 10^5$. The average of these ten thousand transients is considered as the ground truth.

For evaluation, we use a random subset of $1,000$ transients and simulate A-TOFMS using these transients as follows. First, we construct a vector of firing times, $\boldsymbol{\tau} = (\tau_0, \tau_1, \ldots, \tau_{L-1})$ by setting $\tau_0 = 0$ and choosing $\Delta\tau_i \equiv \tau_i - \tau_{i-1}$ uniformly at random from the interval $[\Delta\tau_{\min}, \Delta\tau_{\max}]$. Given $\boldsymbol{\tau}$, the trace is constructed using unaliased transients as prescribed by Eq. (1).

In addition to the aliasing effect, the trace is corrupted by noise. Henceforth, we preprocess the trace before applying the reconstruction algorithm by setting the trace to zero unless it is likely to be the result of an ion impact. As mentioned before, the detector response to each ion impact is a bell shaped pulse which spreads across multiple samples. These pulses are corrupted by electrical noise that can be modeled as additive noise. However, the electrical noise level is significantly smaller in magnitude compared to the detector response to an ion impact and henceforth ion impacts can be marked with high confidence.

We select the potential ion impacts through the following procedure. Define three constants h_0, h_w, and $d_{\min}$ and label a pulse as a potential ion impact event if the width of the pulse at hight h_w is greater than $d_{\min}$. If an event satisfies this criterion the support interval of the event is determined by thresholding the signal at level h_0. Figure 4 demonstrates this procedure through an example. In this figure there are four pulses that exceed the threshold h_w in peak magnitude. From these pulses d_1, d_2, and d_4 satisfy the minimum width condition of $d_i \geq d_{\min}$ at height h_w. Henceforth, after the preprocessing there are three valid events with the start and end times marked at level h_0. We set the trace equal to zero wherever it does not support a valid event. After the preprocessing step, the trace can be represented as a list of events whereby each event describes a single pulse in the trace.

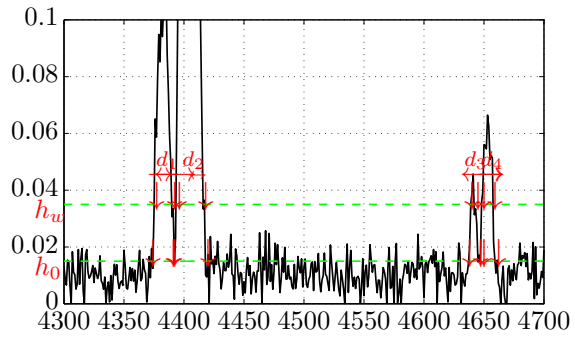

Fig. 1. Preprocessing the data. There are four pulses that exceed the threshold h_w from which three pass the minimum width condition $d_i \geq d_{\min}$ (1, 2, and 4). The markers at the h_0 level indicate the start and end of each marked event.

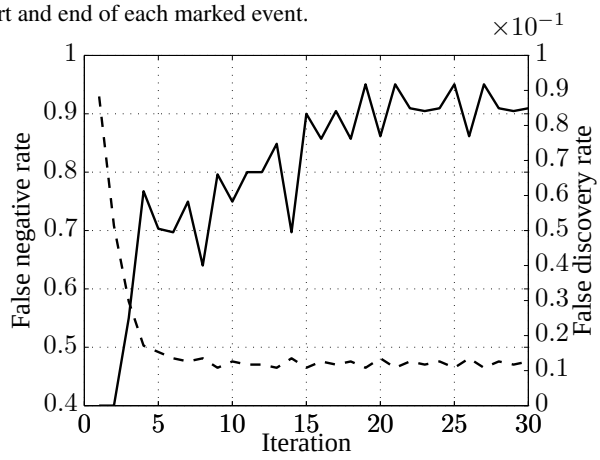

Fig. 2. False negative rate (dashed) and false discovery rate (solid) vs. iteration.

Our procedure to identify valid events also enables us to define metrics for quantitative evaluation of different techniques. We take the true spectrum $\bar{\boldsymbol{x}}$ to be the average of all $10,000$ transients. Let $\widehat{x}$ be an estimate of $\bar{\boldsymbol{x}}$. For some constants h_0, h_w, and $d_{\min}$ define $\bar{\mathcal{E}} = \{\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_{\bar{m}}\}$, and $\widehat{\mathcal{E}} = \{\widehat{e}_1, \widehat{e}_2, \ldots, \widehat{e}_{\widehat{m}}\}$ to be the set of events in $\bar{\boldsymbol{x}}$ and $\widehat{x}$ respectively, obtained through the procedure described above. For two events $\bar{e}_i$ and $\widehat{e}_j$ we say $\widehat{e}_j$ matches $\bar{e}_i$ if $\bar{e}_i$ overlap with at least 50% of the width of $\widehat{e}_j$. For $\bar{e}_i \in \bar{\mathcal{E}}$ we say $\bar{e}_i$ is a false negative if none of the events in $\widehat{\mathcal{E}}$ matches $\bar{e}_i$. For $\widehat{e}_j \in \widehat{\mathcal{E}}$ we say $\widehat{e}_j$ is a true positive if there exist $\bar{e}_i \in \bar{\mathcal{E}}$ such that $\widehat{e}_j$ matches $\bar{e}_i$ and we say $\widehat{e}_j$ is a false positive if it does not match any event in $\bar{\mathcal{E}}$.

Let TP, FP, and FN be the number of true positives, false positives and false negatives respectively. We consider false negative rate (FNR = FN/(TP+FN)), true positive rate (TPR = TP/(TP+FN)), and false discovery rate (FDR = FP/(FP+TP)) as the metrics of interest. (The notion of a true negative is ill-defined in this problem and hence we cannot use the false positive rate metric.)

Unless otherwise stated, the parameters used to obtain the results of this section are as follows. $\mu = 225$, $\gamma = 2.5 \times 10^{-3}$, $\theta_1 =$

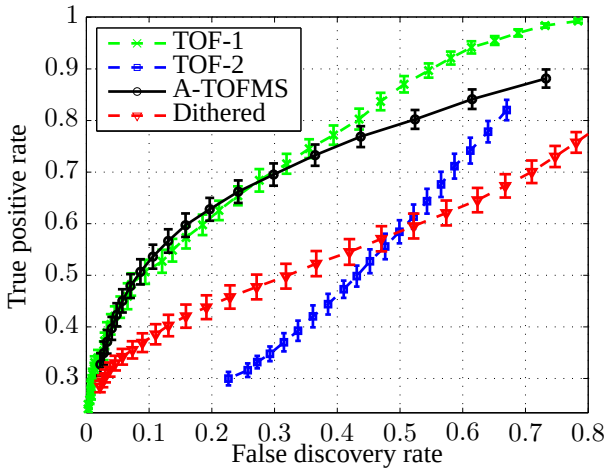

Fig. 3. True positive rate vs false discovery rate

$2\times10^{-2}, \theta_0 = 5\times10^{-4}, h_w^{(\text{trace})} = 2, h_w^{(\text{spectrum})} = 0.2, n = 4\times10^5$, $L = 10^3, \Delta\tau_{\min} = 0, \Delta\tau_{\max} = 2\times10^5$.

Note that $\mathbb{E}\Delta\tau = \frac{1}{4}n$ implies that A-TOFMS is four times more efficient compared to conventional TOFMS in terms of the time it takes to collect the same number of transients. At the same time, each event has on average 4 different position on the spectrum it can be assigned to and the algorithm should be able to infer the correct position with satisfactory accuracy.

Figure 2 shows the false negative rate and false discovery rate as a function of the iteration for a randomsubset of $1,000$ transients and $\theta_0 = 5\times10^{-4}$. The algorithm starts with the all-zero spectrum. Hence, the false negative rate is equal to 1 at iteration 0 and as the algorithm proceeds the false negative rate decreases, converging to a final value of 0.47. The false discovery rate on the other hand increases as the algorithm proceeds settlling at a final value of about 0.085. Inspecting Fig. 2 suggest that the algorithm converges, in the sense of establishing the existence of ions, in about 15 iterations.

To better understand the achieved accuracy, we compare A-TOFMS and three other algorithms in Fig. 3 . The figure shows the true positive rate vs. false discovery rate, as parameterized by θ_0 for the estimated spectrum $\widehat{\boldsymbol{x}}_{A-TOFMS}$ obtained by the A-TOFMS . Each data point is obtained by ten-fold crossvalidation. Namely, we divide the $10,000$ transients into ten buckets of $1,000$ transients each, reconstruct the trace using $1,000$ transients from one bucket and using the average of the other nine buckets as the ground truth. Furthermore, for each bucket we construct 10 random traces by using different random firing times. Error bars indicate $2\hat{\sigma}$ confidence intervals where $\hat{\sigma}$ is the standard error estimate.

We Also plot the corresponding curves for three other algorithms. The red curve corresponds to the *dithered spectrum*, obtained by mapping each event to all possible positions on the spectrum, $\hat{\boldsymbol{x}}_d[i] = \sum_{t\in F}\frac{1}{\deg_t}A_{tj}\boldsymbol{y}[t]$, where A_{tj} is the adjacency matrix defined in Eq. (2) and $\deg_t = \sum_{j=1}^n A_{tj}$ is the number of positions on the spectrum that event t can be mapped to. The dithered spectrum is the most naïve way of processing the trace where one assumes each event to be equally likely to be caused by an ion from any of its potentiall locations on the spectrum.

The green (TOF-1) and blue (TOF-2) curves correspond to conventional TOFMS, i.e., the spectrum obtained by averaging the transients when there is no overlapping, $\widehat{\boldsymbol{x}}_{\text{ave}} = \frac{1}{L}\sum_{l=1}^L \boldsymbol{x}^{(l)}$. In TOF-1, the number of transients is chosen such that the time which takes to perform the TOFMS and A-TOFMS are the same; in this case $L_{\text{TOFMS}} = \frac{1}{4}L_{\text{A-TOFMS}}$ (corresponding to $\mathbb{E}\Delta\tau = \frac{1}{4}n$). TOF-2 corresponds to TOFMS with the same number of transients, which takes 4 times as long to perform compared to the A-TOFMS . Equivalently, one can think of TOF-2 corresponding to the case where an oracle is available for A-TOFMS experiment that to each event assigns its true position on the spectrum.

The dithered and conventional TOFMS curves are parameterized by the thresholding parameter h_w which is used to identify the events in the estimated spectrum. Similar to θ_0, h_w enables us to obtain a trade off between true positive rate and false discovery rate.

This comparison shows that A-TOFMS significantly outperforms the conventional TOFMS. One the one hand, it allows for a four-fold speed-up with essentially unchanged accuracy (comparison with TOF-2). On the other, it allows a two-fold increase in true positive rate for $\text{FDR} = 0.2$ if the experiment duration is unchanged (comparison with TOF-1). It also appears that averaging is not the most effective approach to processing the data, and due to its nonlinear nature A-TOFMS outperforms TOF-1 spectrum at some region of the ROC curve.

5. REFERENCES

[1] W.E. Stephens, "A pulsed mass spectrometer with time dispersion," *Phys. Rev*, vol. 69, no. 691, pp. 46, 1946.

[2] M.T. Roberts, J.P. Dufour, and A.C. Lewis, "Application of comprehensive multidimensional gas chromatography combined with time-of-flight mass spectrometry (GC× GC-TOFMS) for high resolution analysis of hop essential oil," *Journal of separation science*, vol. 27, no. 5-6, pp. 473–478, 2004.

[3] J.B. Fenn, "Electrospray wings for molecular elephants (nobel lecture)," *Angewandte Chemie International Edition*, vol. 42, no. 33, pp. 3871–3894, 2003.

[4] A. Shevchenko, I. Chernushevich, W. Ens, K.G. Standing, B. Thomson, M. Wilm, and M. Mann, "Rapid'de novo'peptide sequencing by a combination of nanoelectrospray, isotopic labeling and a quadrupole/time-of-flight mass spectrometer," *Rapid Communications in Mass Spectrometry*, vol. 11, no. 9, pp. 1015–1024, 1997.

[5] O. Trapp, J.R. Kimmel, O.K. Yoon, I.A. Zuleta, F.M. Fernandez, and R.N. Zare, "Continuous two-channel time-of-flight mass spectrometric detection of electrosprayed ions," *Angewandte Chemie International Edition*, vol. 43, no. 47, pp. 6541–6544, 2004.

[6] F.J. Knorr, M. Ajami, and D.A. Chatfield, "Fourier transform time-of-flight mass spectrometry," *Analytical Chemistry*, vol. 58, no. 4, pp. 690–694, 1986.

[7] A. Brock, N. Rodriguez, and R.N. Zare, "Hadamard transform time-of-flight mass spectrometry," *Analytical Chemistry*, vol. 70, no. 18, pp. 3735–3741, 1998.

[8] G.S. Moore, M. Manlove, and A. Hidalgo, "Statistical Inference Algorithm for DitheredMulti-Pulsing Time-of-Flight Mass Spectrometry," *Submitted*, 2012.

[9] Z. Harmany, R. Marcia, and R. Willett, "This is SPIRAL-TAP: Sparse Poisson Intensity Reconstruction ALgorithms– Theory and Practice," *Image Processing, IEEE Transactions on*, , no. 99, pp. 1–1, 2010.