

MODELLING THE BACKGROUND NOISE IN A MASS SPECTROMETER FOR DENOISING APPLICATIONS

MENTOR: ANTHONY KEARSLEY, NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY,
RICHARD BARNARD, LOUISIANA STATE UNIVERSITY,
YUTHEEKA GADHYAN, UNIVERSITY OF HOUSTON,
YOUZUO LIN, ARIZONA STATE UNIVERSITY,
LIN TONG, IOWA STATE UNIVERSITY,
JIAPING WANG, RUTGERS UNIVERSITY, AND
GUANGJIN ZHONG, MICHIGAN TECHNOLOGICAL UNIVERSITY

1. INTRODUCTION

We look at the problem of denoising a spectrum arising from a mass spectrometer. In order to achieve this, we develop a model which describes the chemical noise using a stochastic differential equation. This model then allows us to use a localized image denoising algorithm to achieve improved denoising. It also is able to predict instrument error and generate additional spectra for simulation purposes. We conclude with some numerical results.

2. MASS SPECTROMETRY

Mass spectrometry is a useful tool in analyzing the chemical makeup of a variety of substances in various fields. For instance, it is useful in the field of combinatorial chemistry, a tool for drug discovery [13]. It is also finding uses in the growing field of microbial forensics [15], an area of increased recent interest [14]. The mass spectrometer in our cases is a matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometer which will give us a measurement of the time of flight of ions that have been charged by a laser. The result is a spectrum which associates intensity with time of flight.

In this paper we consider the problem of denoising data for a given spectrum obtained through MALDI-TOF mass spectrometry performed on a polystyrene sample. A typical spectrum will consist of 25,000 to 100,000 data pairs. One such spectrum can be seen in Figure 2. Close up, the data is clearly highly noisy. In order to analyze the spectrum, peak identification is crucial, as described in [2]. However without

appropriate preprocessing, false peaks may be identified while an inadequate preprocessing algorithm may remove actual peaks, leading to a failure to fully identify the components of the analyte being investigated.

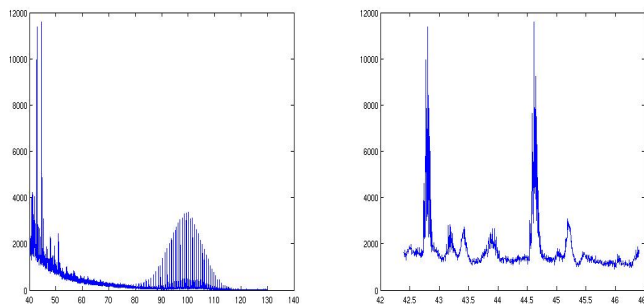


FIGURE 1. A polystyrene’s spectrum, PS1, and a closeup,1000 points shown

Sensitivity in such data arises from two sources: sensitivity due to background ions (chemical noise), and the sensitivity of the instrumentation (instrument noise). As described in [3], sensitivity is currently seen as primarily due to chemical noise as opposed to mechanical or instrument error. Analysis conducted in that article showed “the presence of matrix cluster ions at practically every m/z value.” Thus, methods which can use information on the structure of the chemical noise will more effectively filter out the noise in a sample’s spectrum at multiple points.

Another goal is to reduce operator bias. When considering data analysis methods for spectra arising from spectrophotometry and ligand binding assays, respectively, [5] and [4] dealt with the problems of operator-dependent methods of data analysis. For mass spectrometry, operator-independent methods are crucial, as explained in [6]; specifically they explain that a “uniform method of polymer molecular mass distribution integration is needed to eliminate” wide discrepancies in analysis of polystyrene spectra.

3. THE PROBLEM AND INITIAL ANALYSIS

3.1. Problem Formulation. Given a sample spectrum we will assume that the provided data is the result of additive noise; that is,

$$x_{sample} = x_{true} + \eta$$

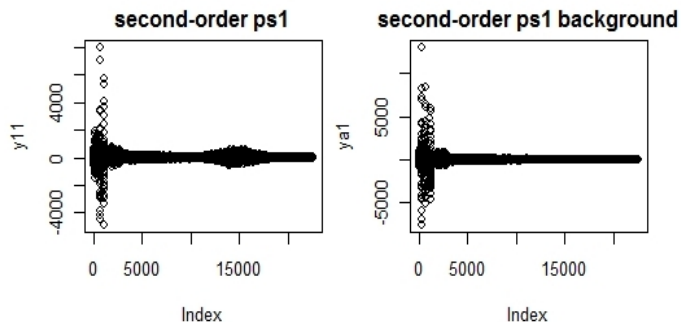


FIGURE 2. PS1 and its background after differencing twice

where x_{sample} is the observed data in a given spectrum arising from a sampled polystyrene, x_{true} is the true signal, and η is a noise term. We are also provided x_{bkgd} which is a spectrum obtained by using the spectrometer with out a sample loaded. The problem we shall address in this paper, then, is to denoise x_{sample} using, in part, information obtained from the background spectrum x_{bkgd} . We will assume in this paper that the noise will be greater when a sample is loaded [1] Under these assumptions, we can filter out a significant portion of the chemical noise from the data, x_{obs} , to obtain a better approximation of x_{true} .

When developing an algorithm to denoise the data we will have to consider several constraints. Our algorithm will need to avoid using any *a priori* knowledge on the noise beyond the above general hypotheses. In light of this and the desire to reduce operator bias, we must also avoid algorithms that involve parameter selections that depend on an operator when analysing x_{obs} ; this will enable us to avoid introducing operator bias. Finally, we wish to ensure that when we denoise our data we are under-denoising our signal. This is a significant requirement, as overly smoothing our data will remove peaks in x_{true} , the locations of which are crucial for analysis of the spectrum.

3.2. An Initial Attempt at Modelling. With the above in mind, we first attempt to model the chemical noise using an ARIMA model. The benefits of such a model, if appropriate, is the relative ease in developing the model, along with an extensive theory having been developed. We first will check stationarity in our background spectrum. If we find stationarity, then we can establish two properties of the chemical noise. First, we will know that its mean and variance are not time-dependent. Secondly, we will see that covariance between points will be a function of how far apart they are in the time only.

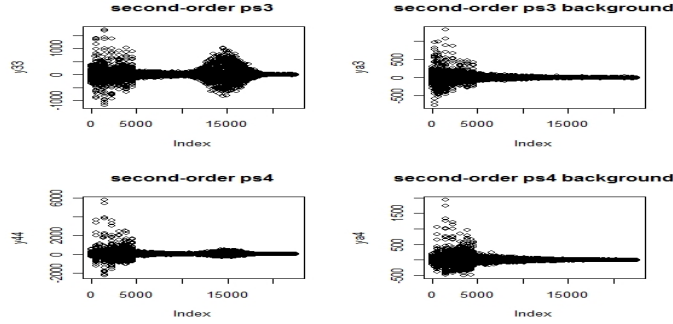


FIGURE 3. PS1, PS3 and their backgrounds after differencing twice

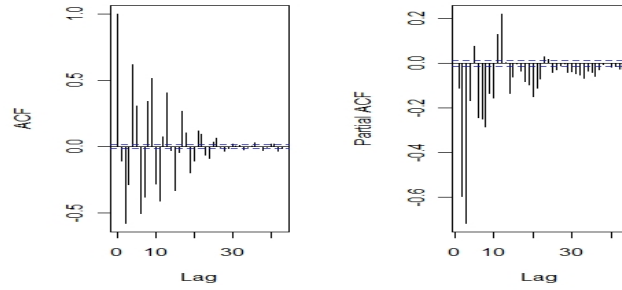


FIGURE 4. ACF and PACF of twice differenced ps1 noise

In both Figure 2 and Figure 3, we see that the second order differencing does not reveal stationarity. However, we notice that the twice differenced PS1 data has an auto-correlation pattern indicating the possibility of seasonality, as seen in Figure 4. Putting aside stationarity considerations for the moment, we attempt to fit the noise to the general seasonal ARIMA model

$$(1) \quad \left(1 - \sum_{i=1}^p \phi_i B^i\right) \left(1 - \sum_{j=1}^P \Phi_j B_s^j\right) W_t = \left(1 - \sum_{i=1}^q \theta_i B^i\right) \left(1 - \sum_{j=1}^Q \Theta_j B_s^j\right) e_t$$

where $B_s W_t = W_{t-s}$, $B_1 = B$ and e_t is white noise. We will take for our modelling $p = P = q = Q = 1$ and $s = 2$. We will need to take $(1 - B)^2 X_t = W - t$ to model the twice differenced data with this fit. We solved for the parameters $\phi, \Phi, \theta, \Theta$ using the standard function in R.

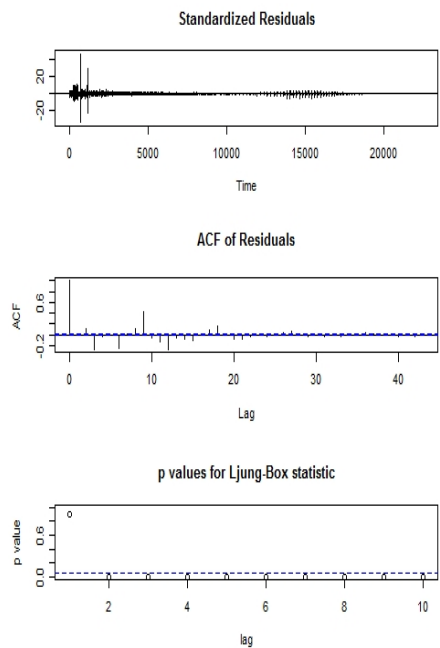


FIGURE 5. Residuals, ACF of residuals, and p-values of Ljung box for Seasonal ARIMA model of PS1 noise

The diagnostic plots shown in Figure 5 reveal that though this seasonal model removes 1-lag auto-correlation successfully, its residual is still severely correlated in other lags. Therefore, time series may not be an appropriate tool here to model our data behavior. This preliminary analysis leads us to consider more intricate models to describe the chemical noise.

4. STOCHASTIC DIFFERENTIAL EQUATION WITH TIME DEPENDENT COEFFICIENTS

Preliminary analysis of the background using time series established non stationarity and time dependent behaviour of the background. We model the dynamics of the background using a general time dependent Stochastic Differential Equation.

$$dX_t = (a_0(t) + a_1(t)X_t)dt + b_0(t)X_t dW_t$$

X_t is the background at time t and $\{W_t\}$ is a Wiener Process, $W_t - W_s \sim N(0, t - s)$, $s < t$

The drift and the diffusion terms in the above equation are both linear functions of X_t . We assume $a_0(t)$ and $a_1(t)$ to be smooth functions of t .

In the next two sections we estimate the parameters and compute the expected value and variance of X_t at each t . We use the local variance of X_t to segment the noisy spectrum and then denoise over each segment separately. The variance over each segment of X_t is computed as the average of the variance over that segment.

4.1. Estimating Parameters. We are given the background data $\{X_{t_i}\}$ at discrete points

$$t_0 < t_1 < \cdots < t_N$$

with $t_{i+1} - t_i = \Delta$.

Using Euler-Maruyama discretization:

$$\Delta X_{t_i} = (a_0(t_i) + a_1(t_i)X_{t_i})\Delta + b_0(t_i)X_{t_i}\Delta W_{t_i}$$

Assuming sufficient smoothness of the coefficients we use the following approximation at any point t_0

$$(2) \quad a_i(t) = a_i(t_0), \quad i = 0, 1$$

for t in a small neighborhood of t_0 . Let h denote the size of the neighborhood. The h is called the bandwidth or window size. We estimate $a_i(t)$ using a locally weighted least squares approximation. We use the following one sided *Epanechnikov Kernel*:

$$(3) \quad K_h(u) = \frac{3}{4h}(1 - u^2) \quad u \in (-1, 0)$$

At each point t_i denoting $a_j(t_i)$ by $a_j(i)$ we solve the following quadratic minimization problem:

$$(4) \quad \min_{a,b} \sum_{j=1}^N \left(\frac{X_{t_{j+1}} - X_{t_j}}{\Delta} - a_0(i) - a_1(i)X_{t_j} \right)^2 K_h\left(\frac{t_i - t_j}{h}\right)$$

$K_h \neq 0$ on $(-1,0)$ and 0 everywhere else, h window of data points considered for regression. We solve the first order conditions to obtain the following expressions for $\hat{a}_j(i)$, $j = 0, 1$ where $\hat{a}_j(i)$, $j = 0, 1$ are the local estimators of $a_j(i)$, $j = 0, 1$.

$$\begin{aligned} \hat{a}_0(i) &= (5) \frac{\sum_{j=1}^N Y_{t_j} K_h(t_j - t_i) - \Delta a_1(i) \sum_{j=1}^N X_{t_j} K_h(t_j - t_i)}{\Delta \sum_{j=1}^N K_h(t_j - t_i)} \\ \hat{a}_1(i) &= \frac{\sum_{j=1}^N K_h(t_j - t_i) \sum_{j=1}^N Y_{t_j} X_{t_j} K_h(t_j - t_i) - \sum_{j=1}^N Y_{t_j} K_h(t_j - t_i) \sum_{j=1}^N X_{t_j} K_h(t_j - t_i)}{\Delta (\sum_{j=1}^N K_h(t_j - t_i) \sum_{j=1}^N K_h X_j^2 - (\sum_{j=1}^N K_h X_j)^2)} \end{aligned}$$

(6)

where $Y_j = X_{j+1} - X_j$.

With these estimates for $a_j(t_i)$ $j = 0, 1$ $i = 1, 2, \dots, N$ we have :

$$\Delta X_{t_i} - (\hat{a}_0(t_i) + \hat{a}_1(t_i)X_{t_i})\Delta \approx b_0(t_i)X_{t_i}\Delta W_{t_i}$$

Using the notation:

$$(7) \quad \frac{\Delta X_{t_i} - (\hat{a}_0(t_i) + \hat{a}_1(t_i)X_{t_i})\Delta}{\Delta} = \hat{E}_i$$

we have from the properties of a Weiner Process, given information up to time t_i

$$b_0(t_i)X_{t_i}\Delta W_{t_i} \sim N(0, b_0(i)^2 X_{t_i}^2 \Delta)$$

Therefore the conditional log-likelihood of \hat{E}_i given X_{t_i} is approximately :

$$\frac{-\log b_0(i)^2 X_{t_i}^2}{2} - \frac{\hat{E}_i^2}{2b_0(i)^2 X_{t_i}^2}$$

In order to estimate $b_0(i)$ we approximate it locally by a constant :

$$(8) \quad b_0(t) = b_0(t_0),$$

in a neighborhood of size h . We maximize the log-likelihood of \hat{E}_i given data in a window of size h . Again we use the *Epanechnikov* one sided kernel to assign weights to the data points according to their distance from the point where b_0 is being estimated. We maximize the following approximate local likelihood function :

$$(9) \quad \max_b -\frac{1}{2} \sum_{j=1}^N K_h\left(\frac{t_j - t_i}{h}\right) \left(\log(b^2 X_{t_i}^2) + \frac{\hat{E}_i^2}{b^2 X_{t_i}^2} \right)$$

Again solving the first order condition for b we have :

$$\hat{b}_0(t_i) = \frac{\sum_{j=1}^N K_h(t_j - t_i) \hat{E}_i^2 |X_{t_i}|^{-2}}{\sum_{j=1}^N K_h(t_j - t_i)}$$

With the estimates for $a_j(i)$ and $b_0(i)$ the dynamics of the background is given by the following discrete approximation of the SDE :

$$\Delta \hat{X}_{t_i} = (\hat{a}_0(t_i) + \hat{a}_1(t_i)\hat{X}_{t_i})\Delta + \hat{b}_0(t_i)\hat{X}_{t_i}\Delta W_{t_i}$$

4.2. Computing the Mean and Variance. Next we compute the mean and variance of X_t at every point t . Going back to the continuous SDE for X_t and observing

$$E \int_{t_0}^t b_0(s) X_s dW_s = 0$$

we have

$$E[X_t] = E[X_{t_0}] + \int_{t_0}^t (a_0(s) + a_1(s)E[X_s]) ds$$

Differentiating both sides with respect to t and denoting $E[X_t]$ by f_t we have the following differential equation:

$$(10) \quad f'_t = a_0(t) + a_1(t)f_t$$

Using a first order forward Euler finite difference scheme with time step Δ we have :

$$\begin{aligned} f_{t_{i+1}} &= f_{t_i} + (a_0(t_i) + a_1(t_i)f_{t_i})\Delta \\ f_{t_0} &= X_{t_0} \end{aligned}$$

Next we compute the conditional variance of $X_{t_{i+1}}$ given information up to time t_i . From the Euler discretization of X_t we have given X_{t_i}

$$\text{Var}(X_{t_{i+1}}) = (b_0(t_i)X_{t_i})^2\Delta$$

Once we have characterized the noise, we turn to denoising the mass spectrum through data preprocessing.

5. DATA PREPROCESSING

The setup of our problem is posed in a canonical inverse model, i.e., by given the measured sample signal, we are going to find the corresponding original clean signal. A regularization technique is often used when either the inverse model is ill-posed or to take into consideration of some kind of prior knowledge of the noise. Two of the more common regularization techniques are Tikhonov Regularization [7] and Total Variation (TV) Regularization [8], where Tikhonov Regularization is formulated as,

$$(11) \quad \min_x \{f(x) = \|Ax - b\|_2^2 + \lambda\|Lx\|_2^2\},$$

and TV Regularization is given by,

$$(12) \quad \min_x \{f(x) = \|Ax - b\|_2^2 + \lambda\|\nabla x\|_1\}.$$

The choice of the technique depends on the characteristics of the problem and our expectations for the solution. TV is well known for its ability to preserve the edges of the image / signal, while Tikhonv is easy to implement from the computational point of view, and there is a much more well-established theory for Tikhonov than TV. In light of this, we first start with Tikhonov Regularization, then briefly discuss the possibility of using TV Regularization in our future work.

5.1. Inverse Model. As we mentioned previously, Tikhonov Regularization (11) will be used in our denoising model. In order to utilize the model, two parameters need to be set before solving for the model, one is the regularization matrix, which is posed as L in (11), and the other one is the regularization parameter, λ in (11). The matrix of L can be the identity, derivative or Laplacian operator which promotes different smoothness across the whole signal; in our numerical test, we used the derivative operator since we tended to see that by using the other two options the reconstructed signal will either be oversmoothed or undersmoothed. When it comes to selecting the other parameter λ , implementing the process becomes more difficult since there is no such straightforward way to pick the appropriate value in the sense of signal to noise ratio (SNR), and the value of λ is very influential on the SNR of the reconstructed signal.

5.2. Parameter Selection. Among the existing parameter selection methods, we can put them all into two categories: those assuming that some *a priori*, knowledge of the noise, such as variance, is given such as, UPRE and discrepancy method [11], and those that do not, such as L-Curve [9], and GCV [10]. In our case, since we can model the background noise—that is, we will be able to know the local variance of the noise—the first type of the parameter estimation methods is appropriate. Specifically we will use UPRE.

UPRE method is short for the Unbiased Predictive Risk Estimator, which is an unbiased estimator of the mean squared error of predictive error P_λ of,

$$(13) \quad \frac{1}{n} \|P_\lambda\|^2 = \frac{1}{n} \|x_\lambda - x_{true}\|^2,$$

where x_λ is the computed solution, and x_{true} is the original clean signal. Since we do not know about x_{true} , Vogel in [11] derives the unbiased estimator for (13) in the Tikhonov case. (For the TV case, please refer

to Lin and Wohlberg's work in [12]) The UPRE formula is given by,

$$\begin{aligned} U(\lambda) &= E\left(\frac{1}{n}\|P_\lambda\|^2\right) \\ &= \frac{1}{n}\|r_\lambda\|^2 + \frac{2\sigma^2}{n}\text{trace}(A_\lambda) - \sigma^2, \end{aligned}$$

where $r_\lambda = x_\lambda - (x_{sample} + \eta)$, $A_\lambda = (I + \lambda I)^{-1}$, σ^2 is the variance of the noise and n is the size of the problem. Based on the above definition, the optimal λ is defined to be,

$$(14) \quad \lambda_{opt} = \min_{\lambda} \{U(\lambda)\}.$$

5.3. Segmentation. Domain decomposition (DD), or segmentation as mentioned in our report, is a technique that people use a lot in dealing with differential equations that having significant different regions in terms of some kind of characteristics. In our problem, one thing is clear: the noise cannot be easily modeled uniformly over the whole region, but we can split the region into several segments and find out the noise characteristics from each one of them, then use them to estimate the corresponding optimal parameters λ on all the segments. Our numerical results shows the strength of using DD by comparing with the results by using one global λ .

6. NUMERICAL TESTS

7. SIMULATING THE SPECTRUM

We fit an SDE to the spectrum in the same manner as the background and obtain the corresponding coefficients. The graph below shows that the simulated data is very close to the actual spectrum. This is significant as generating a spectrum experimentally is very expensive, taking as long as one day. The SDE model can be used to accurately generate spectra. The graph below compares a simulated spectrum to an actual spectrum.

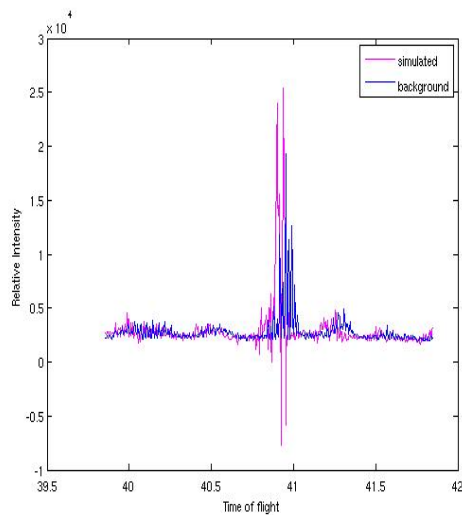


FIGURE 6. Simulated and Actual Spectra

7.1. Picking peaks through variance. The location of genuine peaks in the spectrum can be found by comparing the variance or the error of regression in the background to the spectrum. Low variance regions in spectrum corresponding to high variance in the background can be thought of as fictitious peaks while high variance regions corresponding to low variance in background can be considered to be genuine peaks. Regions where the variance is high both in the background and the spectrum could be a genuine or fictitious peak. This is illustrated in Figure 7 below. The first set of figures correspond to data points between 0 to 500 while the second corresponds to data between 10000 to 11000 points.

7.2. Detecting Instrument Error. The diffusion coefficients $b_0(t)$ that appear in the SDE for the spectrum capture the error in the spectrometer. The quality of the instrument decays slowly at first and then deteriorates rapidly. The plot of b_{t_i} with respect to time captures this expected behavior of the instrument. Figure 8 illustrates the decay in the instrument.

7.3. Generating Synthetic Data. In order to test our algorithm's denoising qualities, we first must create spectra for testing. We used a Savitzky-Golay smoothing process twice on PS1 and PS3, to obtain a smooth spectrum. This was performed with the matlab command `svgolayfit`, using a 2nd order polynomial regression with framesize 45.

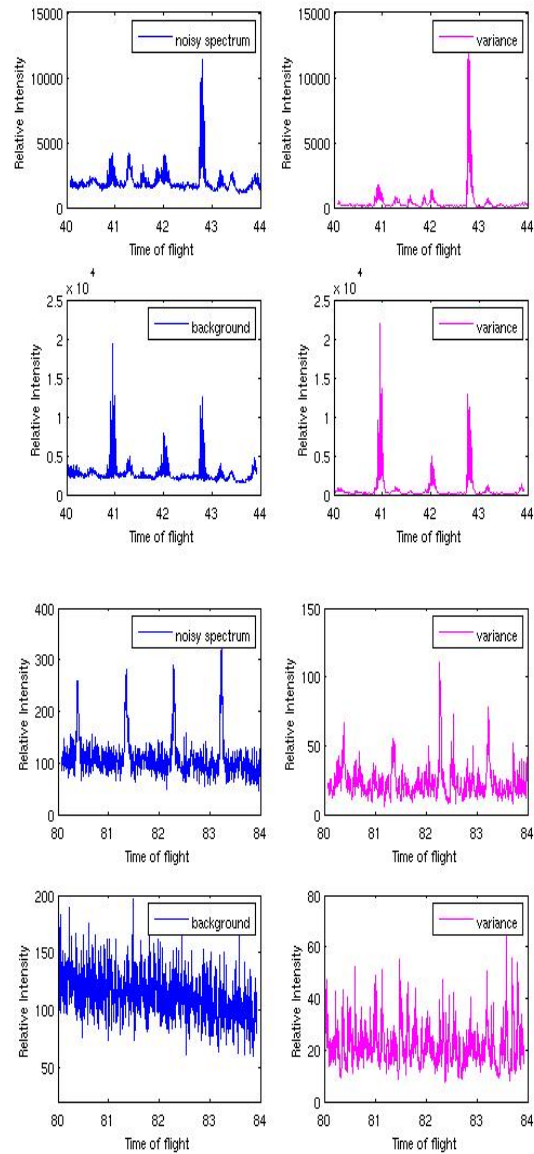


FIGURE 7. Two portions of the spectrum: the first 500 points of the spectrum and points 10000-11000

We selected this method, as opposed to a standard adjacent averaging process, in order to preserve the general structure and locations of local maxima and minima. The resulting smooth data, *smoothedPS1* and *smoothedPS3*, is shown in

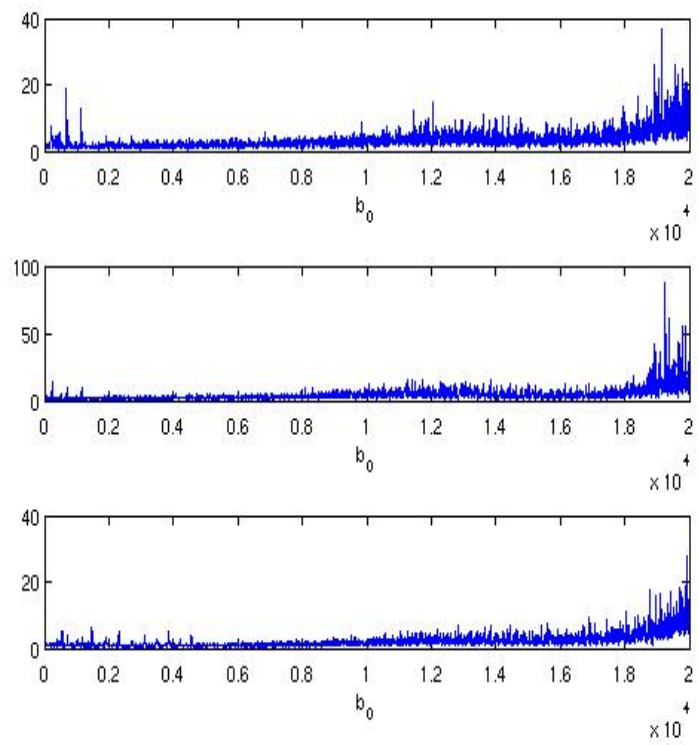


FIGURE 8. Predicting Instrument error

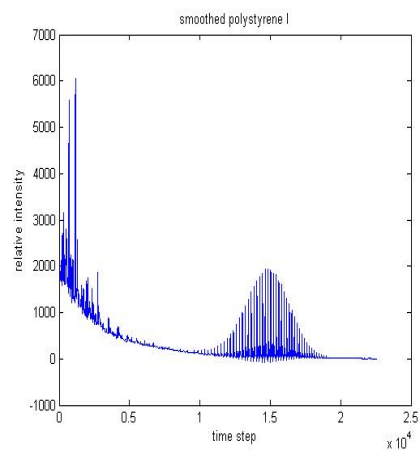


FIGURE 9. Smoothed Version of PS1

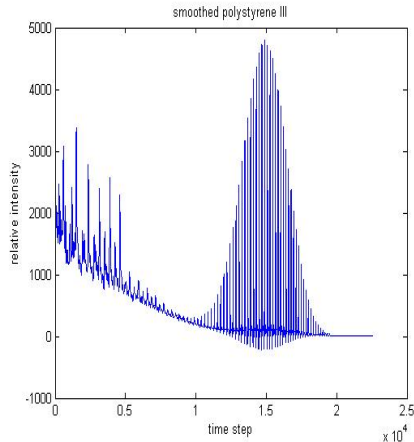


FIGURE 10. Smoothed Version of PS3

We now add noise to *smoothedPS1* and *smoothedPS3* that will mimic the original data. For our purposes, we will generate three test spectra, *test1*, *test2*, from *smoothedPS1* and *test3* *smoothedPS3*. In each example, the noise term, η will be of the form

$$(15) \quad \eta = \alpha \cdot x_{true} \cdot \zeta.$$

For *test1* and *test3* we segment the spectrum at the 10,000th data point, and for *test2* we segment at the 3,000th and 11,000th segments. The term ζ for *test1* is Cauchy on each segment, albeit with different variances for each segment. For *test2* we set ζ to be a Beta term multiplied with a Bernoulli term on each segment, again with differing means and variances on each segment. Finally, for *test3*, ζ is Gaussian on each segment, with different means and variances.

We choose these noise terms to approximate the structure of PS1 and PS2:

Though the noised data sets are close to the original polystyrene spectra, their fundamental structure are different. We can see this from differencing in 13 and 15 or in terms of auto-correlation factor (ACF) in 14 and 16. Thus, though the synthetic data are noisy spectra sharing with a passing resemblance to the actual polystyrene spectra, we can confirm that the background does not resemble a noisy signal originating from one of these simpler sources.

7.4. Denoising Algorithm. Our denoising procedure is as follows. We first use our stochastic differential equation model for the noisy background spectrum to determine the pointwise variances and means

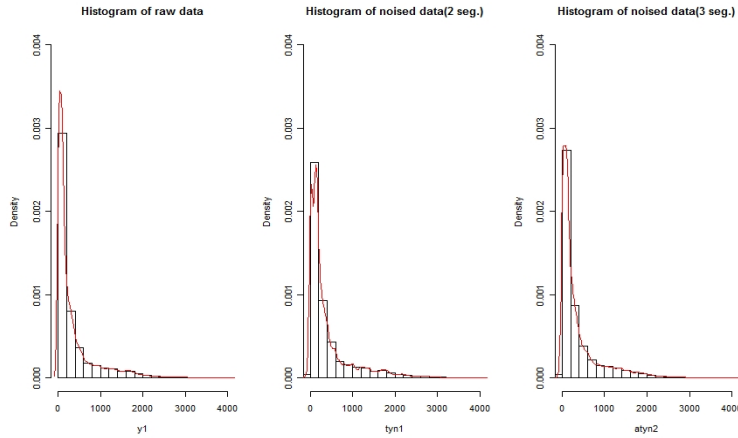


FIGURE 11. Histogram and density plots of, from left to right, PS1, test1, and test3

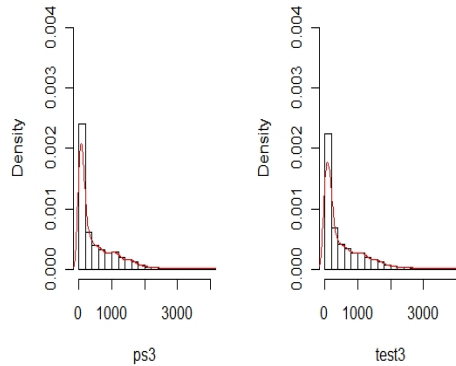


FIGURE 12. Histogram and density plots of PS3 and test2

of the noise. We then partition our spectrum for denoising into segments whose noise behaves in some uniform fashion. In order to do this, we take a rolling integral of the pointwise variance over frames of size 200 points and when we detect a 10% change, we begin a new segment. In order to maintain a minimum image size for our denoising process, we will require a minimum segment length of 1000 data points. After having segmented the data, we then move to denoising the data.

As stated in 5.2, we need to select an appropriate regularization parameter λ in order to denoise our data in an optimal fashion. This will be performed on each segment individually. As stated above, we

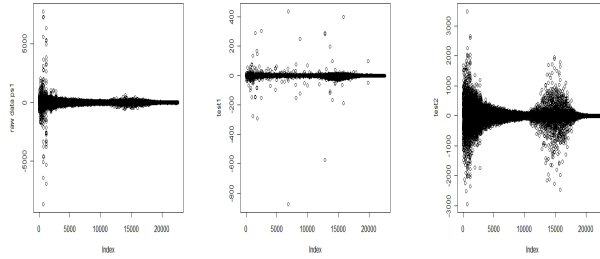


FIGURE 13. Comparison of PS1 with test1,test2 after two differences

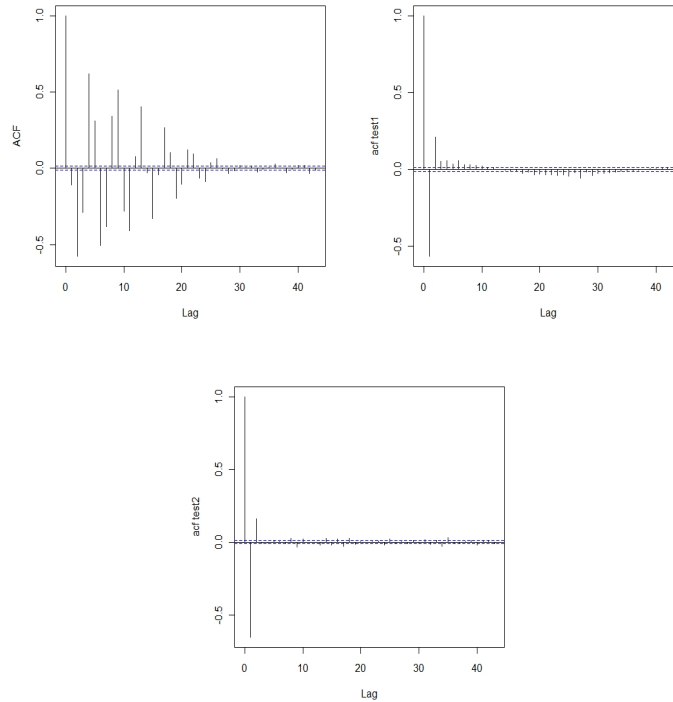


FIGURE 14. Comparison of ACF of PS1,test1,test2

will use UPRE in order find such a λ ; as this algorithm assumes a uniform variance in our noise, we use the mean of the variance on each segment as a way to characterise the noise for our parameter estimation. Having found our parameter, we shift the signal by the pointwise mean of the noise and solve 11 on each segment with that segment's regularization parameter. The shift is due to the requirement

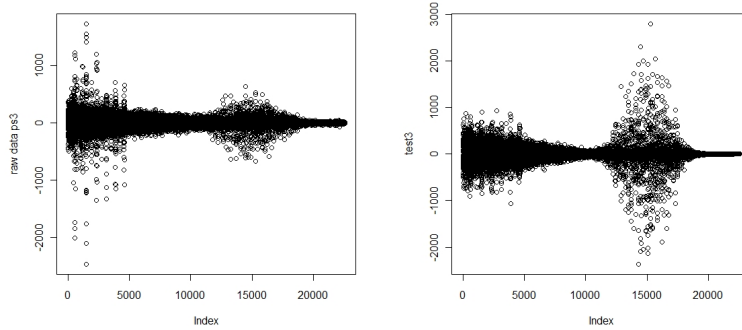


FIGURE 15. Comparison of PS3 with test3 after two differences

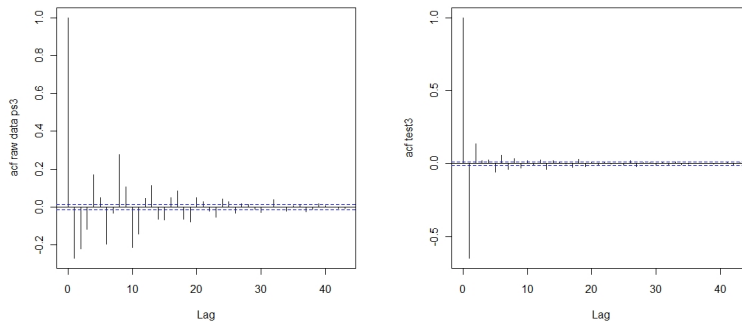


FIGURE 16. Comparison of ACf of PS3, test3

in Tikhonov regularization to have zero-mean noise. This will result in our denoised spectrum.

We performed this algorithm on our test spectra. We took the noise that was added to be the background spectrum. In 7.4 we compare the Signal to Noise Ratio (SNR) for the synthetic data after denoising without segmentation (simply using the variance of the entire spectrum) and by using the above described algorithm. As can be seen, the segmentation delivers a significant improvement in the denoising. The segmentation algorithm is able to better balance data fidelity and smoothing as the variance changes greatly over the entirety of the spectrum.

TABLE 1. SNR of Synthetic Data Before Smoothing, Without Segmentation, and With Segmentation

	Before	Global Tikhonov	Segmented Tikhonov
test1	11.4795	12.4840	19.2765
test2	13.0861	17.9658	19.8942
test3	9.7683	11.4497	15.3959

8. CONCLUSIONS AND FUTURE WORK

We looked at the problem of denoising spectra arising from mass spectrometers. Using background spectra, we modeled the chemical noise that arises in mass spectrometry. We first discarded standard time-series models, such as the ARIMA model, as the noise is highly nonstationary, even after several differencings. We then looked to model the chemical noise using a stochastic differential equation. This model uses only instrument specific-parameters to characterize the noise accurately. We are able to use this model to find the local mean and variance of the chemical noise and simulate additional spectra for future analysis. We then used a locally tuned Tikhonov denoising algorithm to improve denoising over a globally tuned Tikhonov algorithm.

Our segmentation algorithm is currently the only portion of our denoising process involving user-input. We have not investigated the effect of tuning the tolerance in shifts in the tolerance, nor the effect on the minimum segment length. Finally, our segmentation algorithm results in disjoint segments. There may be an improvement in the denoising process if we allow for overlap in our segments and then perform a mortaring procedure. Naturally, the mean of the local variance on each segment may not be the optimal way to characterize the variance on the entire segment, which is a required input for the UPRE algorithm. We would like investigate other criteria for changes in the behavior of the noise that would lead to an automatic segmentation algorithm. We have also not fully investigated other methods for denoising the data, such as TV regularization. However, we expect a similar improvement as TV regularization benefits from a better sense of the noise using the SDE model.

REFERENCES

- [1] Wallace, W. E., Guttman, C. M., Flynn, K. M., Kearsley, A. J., ‘Numerical optimization of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry: Application to synthetic polymer molecular mass distribution

- measurement *Analytica Chimica Acta* Volume: 604 Issue: 1 Special Issue: Pages: 62-68 NOV 26 2007
- [2] Wallace, W. E., Kearsley, A. J., Guttman, C. M., 'An operator-independent approach to mass spectral peak identification and integration *Analytical Chemistry*. Volume: 76 Issue: 9 Pages: 2446-2452. MAY 1 2004
- [3] Kruthinsky, A. N. and Chait, B. T., 'On the Nature of the Chemical Noise in MALDI Mass Spectra' *Journal of the American Society for Mass Spectrometry*. Volume: 13 Issue: 2 Pages: 129-134, 2002
- [4] Little, J. A. 'Comparison of curve fitting models for ligand binding assays' *Chromatographia* Volume: 59 Pages: S177-S181 Supplement: Suppl. SS Published: 2004
- [5] Santana, D. W. E. A., Sepulveda, M. P., Barbeira, P. J. S. 'Spectrophotometric determination of the ASTM color of diesel oil' *Fuel* Volume: 86 Issue: 5-6 Pages: 911-914 MAR-APR 2007
- [6] Guttman, C. M., Wetzels, S. J., Blair, W. M., Fanconi, B. M., Girard, J. E., Goldschmidt, R. J., Wallace, W. E., VanderHart, D. L. 'NIST-Sponsored Interlaboratory Comparison of Polystyrene Molecular Mass Distribution Obtained by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry: Statistical Analysis' *Analytical Chemistry* Volume: 73 Issue: 6 Pages: 152-1262 MAR 1 2001
- [7] A. N. Tikhonov, 'Regularization of Incorrectly Posed Problems' *Soviet Math Dokl.* Volume: 4 Pages: 1624 - 1627, 1963
- [8] L. Rudin and S. Osher and E. Fatemi, 'Nonlinear Total Variation Based Noise Removal' *Physica D*, 60, Pages: 256 - 268, 1992
- [9] P. C. Hansen and D. P. O'Leary, 'The Use of the L-curve in the Regularization of Discrete Ill-Posed Problem' *SIAM Journal on Scientific Computing*. Volume: 14 Pages: 1487 - 1503, 1993
- [10] G. H. Golub and M. T. Heath and G. Wahba, 'Generalized cross-validation as a method for choosing a good ridge parameter', *Technometrics*. Volume: 21 Pages: 215 - 223, 1979
- [11] C. R. Vogel, 'Computational Methods for Inverse Problems' *SIAM*, 2002
- [12] Y. Lin and B. Wohlberg, 'Application of the UPRE Method to Optimal Parameter Selection for Large Scale Regularization Problems' *2008 IEEE Southwest Symposium on Image Analysis and Interpretation* Pages: 89 - 92, 2008
- [13] Kassel, D.B. 'Combinatorial chemistry and mass spectrometry in the 21st century drug discovery laboratory' *Chemical Reviews* Volume: 101 Issue: 2 Pages: 255-267 FEB 2001
- [14] Dance, A. 'Anthrax case ignites new forensics field' *Nature* Volume: 451 Page: 813 AUG 2008
- [15] Jarman, K. H., Kreuzer-Martin, H. W., Wunschel, D. S., Valentine, N. B., Cliff, J. B., Petersen, C. E., Colburn, H. A., Wahl, K. L. 'Bayesian-Integrated Microbial Forensics' *Applied and Environmental Microbiology* Volume: 74, Issue: 11 Pages: 3573-3582 JUNE 2008
- [16] Epanechnikov, V.A. 'Nonparametric estimation of a multidimensional probability density,' *Theor. Prob. Appl.* Volume: 13, Pages: 153-158 1969
- [17] Fan, J., Jiang, J., Zhang, C., Zhou, Z. 'Time-dependent Diffusion Models for Term Structure Dynamics' *Statist. Sinica* Volume: 13 Issue: 4, Pages: 965-992. 2003

- [18] Kloeden, P. E., Platen, E. 'Numerical Solution of Stochastic Differential Equations' Springer, 1999
- [19] Protter, E. 'Stochastic Integration and Differential Equations, Second Edition, Springer, 2005