An Analysis of Noise on Optical Emission Spectroscopy Measurements

Beibei Ma†, Seán McLoone†, John Ringwood† and Niall MacGearailt *

†*Department of Electronic Engineering*
National University of Ireland Maynooth

**Nanomaterials Processing Laboratory*
Dublin City University

E-mail: †beibei.ma@eeng.nuim.ie †sean.mcloone@eeng.nuim.ie

Abstract — **Optical Emission Spectroscopy (OES) is a non-intrusive plasma diagnostic technique that can be used to measure the chemical changes in a plasma that is increasingly being considered for monitoring and control of plasma etch processes. In the practice of collecting OES data for plasma etching, it is inevitable that noise is included in the measurements. The existence of noise can destroy signals or at least make the identification and interpretation of signal patterns unreliable, hence appropriate filtering and pre-processing of the data is needed prior to application of automated feature extraction and analysis techniques. In the absence of *a priori* knowledge of the noise characteristics and system bandwidth the selection of an appropriate noise suppression filter bandwidth is a challenging problem. This paper explores the characteristics of the noise inherent in OES measurements and proposes a systematic method for establishing a suitable noise filter bandwidth based on the auto-correlation and cross-correlation analyses of the filtered signals and their residuals.**

Keywords — **Noise removal, Optical emission spectroscopy, Plasma etch**

I Introduction

Optical Emission Spectroscopy (OES) is a non-intrusive plasma diagnostic technique that can be used to produce rich plasma chemical information in real-time. It operates on the principle that plasmas emit light due to excited electrons within the plasma continually falling from higher to lower energy levels, releasing energy in the form of photons. The wavelengths of these photons are a function of the change in energy levels [1] which are uniquely determined by the structure of chemical species in the plasma. As such, analysis of plasma emission spectra can be used to estimate the instantaneous composition of a plasma and track the density changes of the chemical species overtime.

In modern optical emission spectrometers analysis of light is performed by focusing it onto a diffraction grating using a lens. The grating then redirects the light onto a Charged Coupled Device (CCD) detector with different wavelengths dispersed to different CCD pixels as shown in Fig.1. In the CCD detectors, the photons are detected by a photoactive detection area where they are converted into electrons and accumulated in a capacitor that is periodically discharged for readout.

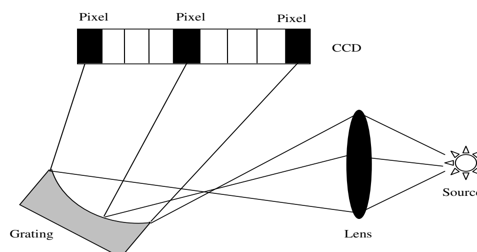

Fig. 1: The diagram of an optical emission spectrometer.

As the semiconductor industry moves towards smaller and smaller feature sizes, OES is increasingly being considered for process monitoring and control of plasma etch processes [2], a key step in the manufacture of modern integrated circuits. In a typical plasma etch chamber (Fig. 2) gas is pumped into the chamber under vacuum

and ionised using a high power Microwave (MF) source to create a plasma. A radio frequency (RF) electromagnetic field accelerates the resulting ionised species towards the electrode, where they interact both chemically and physically with the wafer, etching away the exposed surface. OES data, which consists of measurements of the optical emission intensities as a function of wavelength and time, is generally collected from the exhaust plasma leaving the chamber using a multichannel spectrometer. As an example, Fig. 3 shows the OES data obtained for a typical polysilicon wafer etch step collected using an Ocean Optics USB2000 spectrometer with a CCD detector consisting of 2045 pixels (corresponding to 2045 wavelengths, ranging from 170nm to 875nm) with a sampling interval of 0.75s.

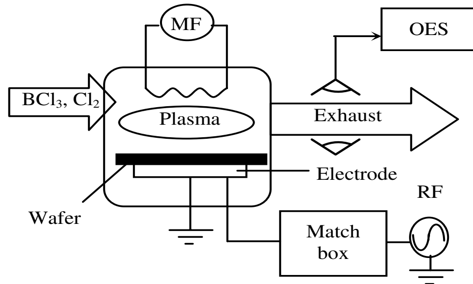

Fig. 2: A diagram showing the basic features of a plasma etching chamber: MF=Microwave Frequency; RF=Radio Frequency.

OES data is characterized by high dimensionality and substantial redundancy making direct visualization and recognition of useful features and key wavelengths difficult. Manual inspection is in general impractical due to the huge volumes of data involved in a typical manufacturing scenario, requiring robust and automated data reduction, feature extraction and analysis techniques. As a precursor to the application of such techniques a proper treatment of the noise on the OES signals is needed, that is, channels with low signal-to-noise ratio need to be removed and appropriate filtering applied to the remainder.

It is inevitable that OES signals contain noise. When significant noise is present the signals can be completely destroyed or, at the very least, obscured to the extent that the identification and interpretation of signal patterns becomes unreliable. Hence, it is necessary to produce an estimate of noise level so that appropriate pre-processing and filtering can be applied to clean up the signals for better pattern extraction.

A number of possibilities exist for estimating the noise level on OES signals. If it is feasible to conduct controlled experiments in-situ a direct estimate of the background noise level can be obtained by collecting OES data in the absence of a plasma or with the OES CCD detector occluded. However, this approach is not reliable as OES measurement noise is known to have signal dependent characteristics and can vary over time, for example, due to changes in the absorption characteristics of the plasma chamber wall, chamber temperature and plasma gases.

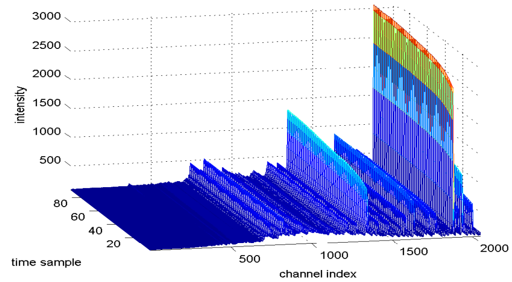

Fig. 3: Plasma etch OES data for a single wafer, recorded over a complete etch step.

An alternative approach is to estimate the noise level from inactive OES channels during normal operation, that is, channels that are recording signals at wavelengths where the plasma is not emitting light. In principle the inactive channels can be established *a priori* from knowledge of the plasma chemistry and excitation levels, but in practice the complex chemistry and correspondingly dense spectral emissions of the plasmas employed in etch processes makes this infeasible. Consequently, inactive channels are generally selected manually by empirical observation which is clearly subjective and can be unreliable since low intensity emissions can be difficult to detect in the noise. It should also be noted that using inactive channels as a noise reference cannot take account of signal dependent noise characteristics.

A third approach which overcomes the issues with the previous methods is to estimate the noise level on each channel following application of an appropriately selected noise suppression filter. Defining the raw signal of the ith OES channel as $\mathbf{x}_i$ and the filtered signal as $\mathbf{x}_i^f$, the noise power can be estimated as

$$N_i = pow(\mathbf{x}_i - \mathbf{x}_i^f) = E[(\mathbf{x}_i - \mathbf{x}_i^f)^2], \quad (1)$$

where $pow(\cdot)$ denotes the signal power and $E[\cdot]$ is the expectation operator. Furthermore, the signal-to-noise ratio (SNR), the standard method for measuring the strength of a signal relative to the noise, can be estimated as

$$SNR_i = pow(\mathbf{x}_i^f)/pow(\mathbf{x}_i - \mathbf{x}_i^f). \quad (2)$$

The validity of this approach relies on the user having predetermined the appropriate bandwidth

for the noise suppression filter. In the absence of *a priori* knowledge of the characteristics of the noise and OES signal bandwidth this is a challenging problem, and the focus of the remainder of this paper.

In Section two the main characteristics of OES noise are described. Section 3 considers the correlation that exists between channels and the use of PCA to detect its presence. Local correlation analysis is proposed to estimate the number of channels that are highly correlated. Then a systematic method for establishing a suitable noise filter bandwidth is presented based on the auto-correlation and cross-correlation analyses of the filtered signals and their residuals. Throughout the paper the example data set presented in Fig. 3 is used as a case study.

II Noise Sources

Before the noise analysis is presented, it is beneficial to have an awareness of the different noise generation sources in OES. Normally, noise is considered as being the high frequency variations in signals. However, in plasma etch OES data, not all the variations are caused by noise. As such, we divide different variation sources into two categories: process variation and sensor noise, where the process variation refers to any variations occurring in the plasma emission system and sensor noise refers to the noise occurring in the plasma measuring system.

Some process variations are driven by recipe changes, fluctuations in control variables, external disturbances and the changes in the plasma chemistry as a result of etching byproducts, but these are all relatively low frequency variations. The high frequency process variations are driven by the stochastic nature of the plasma emission process and the inherent instability of the volatile etch by-products generated by the plasma at room temperature. The interactions between the molecules and atoms of different etch by-products are unpredictable, causing great uncertainty in the chemical optical emissions.

The principal noise sources in the measurement system are: shot noise, caused by the random fluctuations in photon arrival times; thermal noise (or dark noise), generated by thermal agitation of electrons, the electronic components and wire; and readout noise, including the conversion from an analogue signal to a digital number. The strength of shot noise increases with signal strength and is generally the dominant noise for large signals.

From the list of sensor noise sources, we can see that most noise can be regarded as white noise (broadband noise). Low pass filtering of the signal can be used to suppress the noise, but only at the expense of also suppressing the contributions of the high frequency process variations. However, in practice this is not a concern as for monitoring and control purposes we are only interested in patterns generated by recipe changes, phase transitions and process transients that operate at a much lower frequency.

III Selecting the Filter Bandwidth

In using filtering to remove the high frequency noise from the OES case study data set (Fig. 3), a low-pass Butterworth filter is used in preference to other widely used filters such as the Chebyshev Type I/Type II filter and the elliptic filter, as it is characterized by having a flat gain in its pass band and hence, provides minimal distortion of the filtered signals. Due to the slow roll-off into the stop band, a high-order Butterworth filter is needed to obtain faster roll-off. Here we employ a 4th-order Butterworth filter.

a) *Selecting the Filter Bandwidth Based on Single Channels*

To determine the cut-off frequency of the low-pass filter, we begin by employing the DFT (Discrete Fourier Transform) [3] to view the signal frequency distribution. Taking the strongest OES signal as an example, Fig. 4 (a) and (b) show the DFT and the cumulative PSD (Power Spectral Density) analysis of the signal, respectively. The cumulative PSD analysis shows that 76.3% of the signal power is contained in the bandwidth 0-0.1Hz and 88.2% of the signal power contained in the bandwidth 0-0.3Hz.

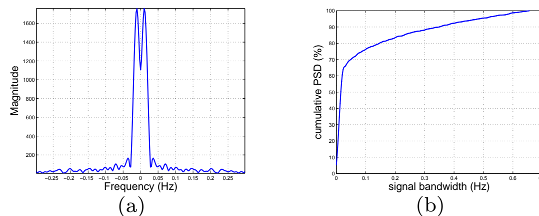

Fig. 4: Analysis of the strongest signal: (a) DFT analysis; (b) Cumulative PSD analysis

Let $\mathbf{x}_i(\mathbb{R}^{m\times 1})$ denote the raw signal from the ith OES channel, for a given low-pass Butterworth filter, the filtered signal ($\mathbf{x}_i^{\text{f}}$) is defined as

$$\mathbf{x}_i^{\text{f}} = filt(\mathbf{x}_i) \tag{3}$$

where $filt(\cdot)$ is the butterworth filtering function and the residual $\mathbf{x}_i^{\text{r}}$ is given by

$$\mathbf{x}_i^{\text{r}} = \mathbf{x}_i - \mathbf{x}_i^{\text{f}}. \tag{4}$$

One approach to selecting the best low-pass filter (LPF) bandwidth, f^{B}, is to choose the frequency

that minimises the correlation between the filtered signal ($\mathbf{x}_i^{\mathrm{f}}$) and the residual signal ($\mathbf{x}_i^{\mathrm{r}}$), *i.e.*

$$f^{\mathrm{B}} = \arg\min_{f} |corr(\mathbf{x}_i^{\mathrm{f}}, \mathbf{x}_i^{\mathrm{r}})|, \quad (5)$$

where $corr(\cdot)$ denotes the correlation function. A correlation analysis of the filtered signal and residual signal, as a function of LPF bandwidth is shown in Fig. 5 for the two selected signals.

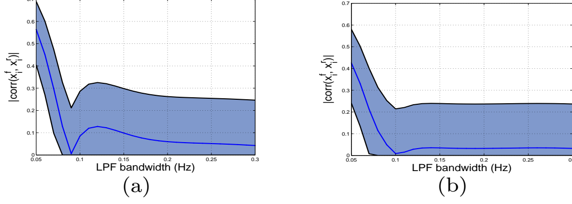

Fig. 5: Correlation between the filtered signal and residual signal as a function of f (shaded area showing the 95% confidence interval for the correlation coefficient estimates): (a) Strongest signal; (b) The signal with the third highest power

In Fig. 5 the shaded area shows the 95% confidence interval for the correlation coefficient estimates. Fig. 5 (a) shows that the minimal correlation is obtained when $f = 0.09$Hz and that the correlation is statistically insignificant for $f \geq 0.0755$Hz. Using the same method, $f^{\mathrm{B}} \geq 0.0708$Hz for the second signal.

To evaluate the consistency of this approach, f^{B} was computed for each of the OES signals. Based on a DFT and cumulative PSD analysis, the search range for f^{B} was set between 0.001Hz and 0.2Hz. A plot of f^{B} for all the OES channels is shown in Fig. 6 (a), which shows that there is significant uncertainty in the value of f^{B} across channels and that there is some local correlation in values. Fig. 6 (b) reveals that the value of f^{B} varies as a function of signal power with the value of f^{B} at which correlation becomes insignificant, decreasing as the power in the signal decreases. In addition the spread in f^{B} increases substantially as signal power decreases (from 0.06-0.08Hz for high power signals to 0.01-0.12Hz for low power signals).

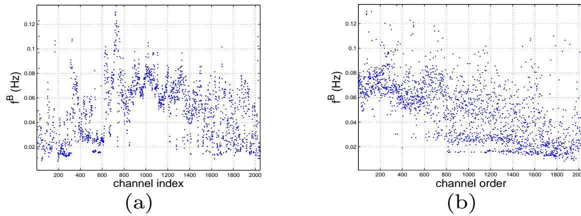

Fig. 6: f^{B} for each OES signal: (a) Channels ordered sequentially; (b) Channels in order of decreasing signal strength.

Since f^{B} estimates in Fig. 6 (b) are lower bounds on suitable LPF bandwidths and the values computed for the largest signal powers are the most reliable, $f^{\mathrm{B}} \geq 0.09$Hz represents a good compromise (96.2 % of channels) and choosing $f^{\mathrm{B}} = 0.1$Hz achieves minimal correlation for most channels (98.1 % of channels).

b) PCA of the Residual Signals

The single channel analysis in the previous subsection suggests that an appropriate LPF bandwidth is 0.1Hz. In this section, Principal Component Analysis (PCA) is used to look at the patterns contained in the residuals across channels. The residuals obtained from filtering all channels with a 0.1Hz bandwidth LPF are shown in Fig. 7 (a). A PCA analysis of the residuals reveals that the first PC explains 12.61% of the variance, with 4.48% and 3.98% variance explained by the second and third PCs, respectively. The variance explained by the first 20 PCs is shown in Fig. 7 (b). This confirms that inter-channel patterns exist in the residual data (*i.e.* some of the residual signals are correlated). If the residual signals were independent and identically distributed random noise, then the variance explained by each PC will be nearly the same. Applying PCA to the residual signals obtained from filtering using various filter bandwidths shows that inter-channel correlation still exists in the residuals. This suggests that there may be an intrinsic correlation in the signals. In fact, because of the limited spectral resolution of OES spectroscopes, the optical emission at a given wavelength will be detected over a number of adjacent channels, leading to local correlation. In the next subsection, an analysis of the extent of this local correlation is given. Note that the existence of local correlation in the residuals corroborates the existence of high frequency process variations.

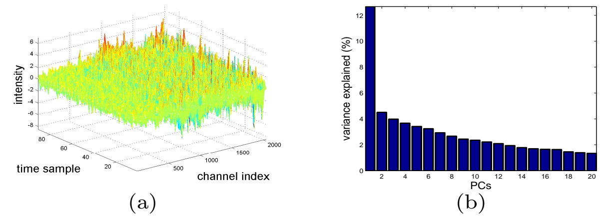

Fig. 7: (a) Residual signals, obtained by subtracting the filtered signals from the original signals (LPF bandwidth = 0.1Hz); (b) Variance explained by each PC in a PCA analysis of the residual signals.

c) Local Correlation

Due to the limited OES spectroscope resolution, the optical emission signals are detected simulta-

neously by a number of adjacent OES channels. To assess the extent of this spectral spread, the local correlation between channels is investigated. The local correlation, r_j^L, is defined as the correlation between the signal ($\mathbf{x}_i$) from channel i and the signal from the channel, j channels away from channel i, ($i.e.$ $\mathbf{x}_{i+j}$) averaged over all channel positions, that is:

$$r_j^L = \frac{1}{n-j}\sum_{i=1}^{n-j} corr(\mathbf{x}_i^r, \mathbf{x}_{i+j}^r), \text{ for } j \geq 0 \quad (6)$$

$$= \frac{1}{n+j}\sum_{i=1-j}^{n} corr(\mathbf{x}_i^r, \mathbf{x}_{i+j}^r), \text{ for } j \leq 0. \quad (7)$$

A plot of r_j^L for the 0.1Hz LPF residual is given in Fig. 8. It can be seen that the spectral spread is ±5 on either side of a given channel. The shaded area shown in Fig. 8 denotes the 95% confidence interval associated with each r_j^L.

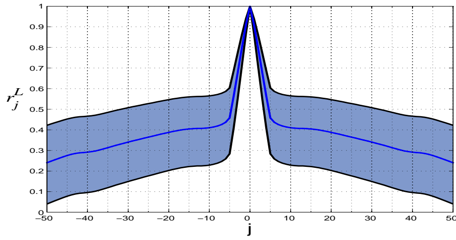

Fig. 8: Local correlation, r_j^L, for the 0.1Hz LPF residual, $\mathbf{x}^r$.

d) Crosscorrelation of the Residual Signals

To estimate the inter-channel correlation between the residual signals, the following average crosscorrelation is used:

$$\bar{r}^C = \frac{1}{n-l}\sum_{i=1}^{n-l}\left(\frac{\sum_{j=i+l}^{n} corr(\mathbf{x}_i^r, \mathbf{x}_j^r)}{n-i-l+1}\right), \quad (8)$$

where $\mathbf{x}_i^r$ and $\mathbf{x}_j^r$ are the ith and j^{th} residual signals, respectively and l denotes the number of adjacent channels that are correlated due to spectral leakage. From the analysis in Section c), $l = 6$. By omitting the locally correlated channels, the value of $\bar{r}^C$ better reflects the filter bandwidth dependent correlation in the residual signals. Fig. 9 shows a plot of $\bar{r}^C$ as a function of LPF bandwidth and the shaded area denotes the 95% confidence interval. The value of $\bar{r}^C$ decreases with increasing LPF bandwidth and drops below 0.1 for $f^{\text{B}} \geq 0.05$Hz and 0.05 for $f^{\text{B}} \geq 0.1$Hz.

e) Crosscorrelation between the Residual Signals and Filtered Signals

Crosscorrelation analysis is used to estimate the correlation between the filtered signals and residuals. Denoting $\bar{r}_i$ as the correlation between $\mathbf{x}_i^{\text{f}}$ and $\mathbf{x}_i^{\text{r}}$ (as defined in Eq. (3) and Eq. (4), respectively), the averaged correlation coefficient ($\bar{r}$) for all the channels is defined as

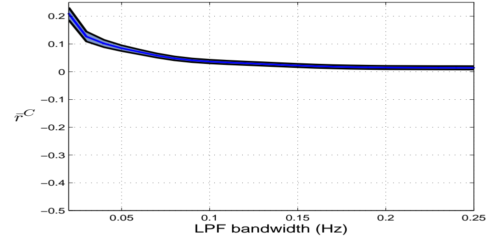

Fig. 9: Changes of $\bar{r}^C$ as a function of LPF bandwidth.

$$\bar{r} = \frac{1}{n}\sum_{i=1}^{n}\bar{r}_i, \quad (9)$$

where $\bar{r}_i = corr(\mathbf{x}_i^{\text{f}}, \mathbf{x}_i^{\text{r}})$ and n is the number of channels. Fig. 10 shows the change of $\bar{r}$ as a function of LPF bandwidth. It can be seen that the averaged correlation between the filtered signals and residuals is insignificant for $f^{\text{B}} \geq 0.1$Hz.

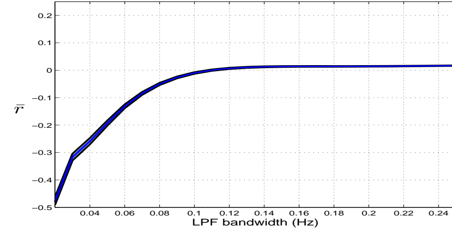

Fig. 10: Changes of $\bar{r}$ for different LPF bandwidths.

f) Autocorrelation of the Residual Signals

Autocorrelation analysis is employed to estimate the correlation between the residual signal and its time-lagged values. If the residual signal is noise, then no significant correlation should exist. To measure the autocorrelation for all the residual signals, we define a new function, $\bar{r}^A$:

$$\bar{r}^A = \frac{1}{2(m-1)n}\sum_{j=-m+1, j\neq 0}^{m-1}\left|\sum_{i=1}^{n} corr(\mathbf{x}_i^r, \mathbf{x}_i^r(j))\right|, \quad (10)$$

where n is the number of OES channels and $\mathbf{x}_i^r(j)$ is the jth lagged signal of $\mathbf{x}_i^r$, ($\mathbf{x}_i^r \in \mathbb{R}^{m\times 1}$). Hence, $\bar{r}^A$ measures the averaged correlation levels between the residual signal and its lagged signals over all channels. The smaller the value of $\bar{r}^A$, the lower the correlation in the residuals. Fig. 11 shows the variation in $\bar{r}^A$ as a function of LPF bandwidth and the shaded area denotes the 95% confidence interval. The value of $\bar{r}^A$ decreases with increasing LPF bandwidth and drops below 0.05 for $f^{\text{B}} \geq 0.05$Hz and close to 0 for $f^{\text{B}} \geq 0.1$Hz.

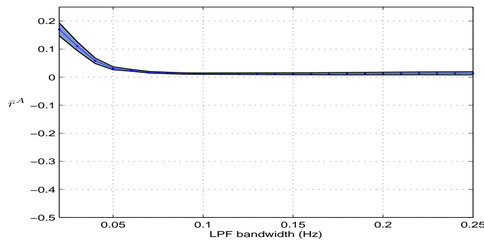

Fig. 11: Changes of $\bar{r}^A$ for different LPF bandwidths.

g) Selection of the LPF Bandwidth

While the analysis in the previous sections cannot provide an exact optimal solution to the LPF bandwidth, it is clear that there is no method that can be self-sufficient. As shown in Table 1, the lower bounds on f^{B} identified using the different techniques are relatively consistent. Thus, for the case study considered, 0.1Hz represents a balanced choice for the noise suppression filter bandwidth.

Method	LPF Bandwidth, f^B, (Hz)
Single signal based	$f^B \geq 0.09$
$\bar{r}^C$	$f^B \geq 0.1$
$\bar{r}$	$f^B \geq 0.1$
$\bar{r}^A$	$f^B \geq 0.05$

Table 1: LPF bandwidth selected by different methods

IV Filtering Result Visualization

The results of filtering the raw OES signals using a 0.1Hz, 4$^{\text{th}}$ order Butterworth LPF are shown in Fig. 12. Plot (a) shows the filtered signals and plot (b) the residuals, respectively.

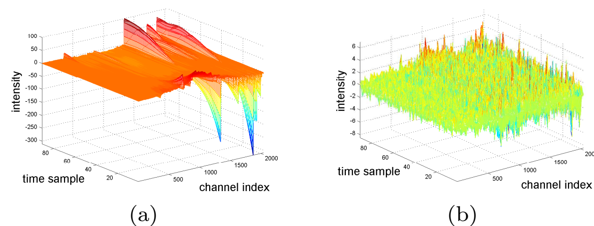

Fig. 12: The raw OES data filtered by a 4$^{\text{th}}$ order low-pass Butterworth filter with cut-off frequency set at 0.1Hz: (a) Filtered signals; (b) Residuals

To observe the effect of filtering on individual channels, the results for three channels are shown in Fig. 13. These channels correspond to the 33.33%, 66.67% and 98.85% division points of the cumulative signal power plot (sorted in descending order) as shown in Fig. 13 (a). This also confirms that the selection of 0.1Hz as the LPF bandwidth is reasonable for filtering the OES signals.

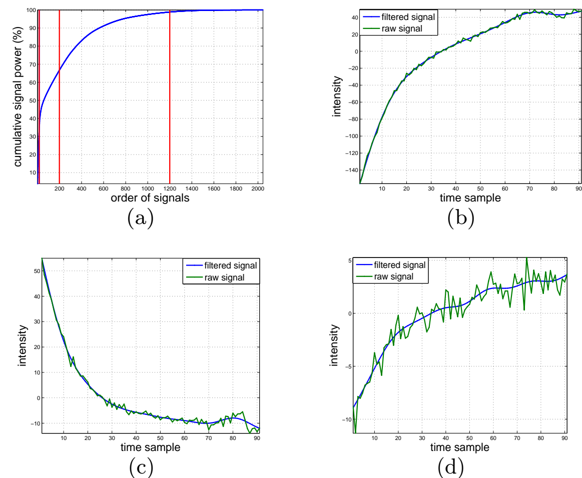

Fig. 13: Filtering results for three selected signals: (a) Cumulative signal power plot for signals sorted in descending power order; (b), (c) and (d) show the filtering results for the signals corresponding to the 33.33%, 66.67% and 98.85% division points of the cumulative signal power plot

V Conclusions

In this paper, an analysis has been presented of the noise in OES data from plasma etch processes. Using a representative case study, it has been highlighted that the noise varies as a function of signal power/wavelengths and that noise on successive channels is correlated due to the limited spectral resolution of spectrometers. Experimental results show that the proposed systematic method of using auto-correlation and cross-correlation to analyse the filtered and residual signals can provide an effective solution to the problem of selecting the filter bandwidth for OES signal filtering and noise removal.

Acknowledgements

The authors gratefully acknowledge the financial support of Enterprise Ireland (grant IP/2006/0325) and Intel (Ireland) Ltd.

References

[1] M. Sugawara. *Plasma Etching: Fundamentals and Applications*. Oxford University Press, New York, 1998.

[2] M. Splichal and H. Anderson. Application of chemometrics to optical emission spectroscopy for plasma monitoring. *Proceedings of SPIE*, 2:189–203, 1987.

[3] E. Brigham. *The fast fourier transform and its applications*. Prentice Hall, 1988.