Journal of Magnetic Resonance 195 (2008) 17–22

Contents lists available at ScienceDirect

Journal of Magnetic Resonance

journal homepage: www.elsevier.com/locate/jmr

Real time exponentially weighted recursive least squares adaptive signal averaging for enhancing the sensitivity of continuous wave magnetic resonance

C.J. Cochrane, P.M. Lenahan *

The Pennsylvania State University, 101 EES Building, University Park, PA 16802, USA

ARTICLE INFO

Article history:
Received 14 March 2008
Revised 8 August 2008
Available online 20 August 2008

Keywords:
Signal processing
Semiconductor reliability
Electrically detected magnetic resonance
Signal averaging
Signal to noise

ABSTRACT

This study involves the use of adaptive signal processing techniques to improve the sensitivity of continuous wave electrically detected magnetic resonance. The approach should be of widespread utility in continuous wave magnetic resonance experiments of all kinds. We utilize adaptive signal averaging to expedite the averaging process usually performed in magnetic resonance experiments. We were capable of reducing the noise variance in a single trace by a factor of 11.3 which is equivalent to reduction in time by the same factor. This factor can be quite significant especially when signal averaging must be performed over the span of many hours to days. This technique may also be tailored to conventional electron spin resonance experiments and other techniques where signal averaging is utilized. The approach may offer promise in the eventual development of spin based quantum computing.

1. Introduction

This paper describes a real time exponentially weighted recursive least squares adaptive signal averaging technique which greatly decreases the amount of time needed for signal averaging of continuous wave magnetic resonance measurements. The technique provides a very low cost means to achieve a quite significant improvement in signal to noise ratio and data acquisition time. The technique is utilized in electrically detected magnetic resonance (EDMR) via spin dependent recombination (SDR) [1,2] in individual transistors [3,4]. However, the approach should be widely applicable in continuous wave magnetic resonance measurements.

EDMR typically involves SDR. EDMR in general and SDR in particular are electron spin resonance (ESR) techniques [5] in which a spin dependent change in current provides a very sensitive measurement of paramagnetic defects. Without special application of digital signal processing techniques, EDMR measurements involving SDR are about seven orders of magnitude more sensitive than conventional ESR. The techniques are therefore particularly useful in studies of imperfections in the semiconductor devices utilized in integrated circuits. In such devices, the dimensions are quite small and can have very low defect densities. SDR detected EDMR can be utilized in fully processed devices such as metal oxide semiconductor field effect transistors (MOSFETs), bipolar junction transistors (BJTs), and diodes [3,4]. With some additional improvements, the technique's very high sensitivity may make it potentially useful for single spin detection and quantum computing. However, the sensitivity EDMR is not currently high enough to detect a single spin in the presence of the noise encountered with present day EDMR spectrometers in a reasonable amount of time. The purpose of this study is to improve the signal to noise ratio and thus the rate of data acquisition in such sensitive EDMR measurements.

Continuous wave magnetic resonance typically utilizes a sinusoidal modulation of the applied magnetic field, thereby encoding the signal in a sinusoid. The amplitude of the modulated signal is a measure of the EDMR and therefore, a measure of the number of defects within the device. A lock-in amplifier (LIA) is then used to demodulate the amplitude modulated EDMR signal to DC, thus exploiting the sensitivity enhancement available from the phase and frequency detection. This widely used method effectively attenuates much of the noise associated with the 1/f noise typically observed with a DC current produced by the diode or transistor utilized in the EDMR measurement. However, lock-in detection alone is not always sufficient to achieve a reasonable signal-to-noise ratio (SNR), so signal averaging is also often utilized. In very small devices with very low defect densities, the EDMR signal to noise ratios can be low. Therefore, extensive signal averaging may be required to achieve a reasonable SNR for the EDMR spectra.

Though work has been performed to remove noise observed in related fields via software, such as nuclear magnetic resonance (NMR) [8], not much has been done in any area of ESR including EDMR. There are many things can be done to improve EDMR SNR in terms of hardware such as proper grounding, minimizing cable length to reduce stray capacitance, and utilizing low noise preamplifiers. However, the focus of this paper will be on the advancements that we made in software.

* Corresponding author. Fax: +1 814 863 7967.
E-mail addresses: corey.cochrane@gmail.com (C.J. Cochrane), pmlesm@engr.psu.edu (P.M. Lenahan).

1090-7807/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2008.08.004

We performed our measurements on 4H SiC lateral n-channel MOSFETs fabricated by Cree Corporation. These devices had a gate area of 200 × 200 μm^2 and a thickness of 500 Å. These devices received a thermal ONO gate growth process. All EDMR measurements were made with the sample at room temperature and were performed with a fixed gate voltage. All EDMR spectra reported here were taken with the magnetic field orientation parallel to the SiC/oxide surface normal which is nearly parallel to the crystal's *c*-axis. EDMR measurements were made with a modulation frequency of 1400 Hz and quite low modulation filed amplitude (<0.1 G). The EDMR measurements were made on a custom built EDMR spectrometer which utilizes a Resonance Instruments 8330 X-band bridge, TE_{102} cavity, and magnetic field controller, a Varian E-line century 4 in. magnet, and power supply. We use a Stanford Research Systems SR570 current preamplifier to prefilter and amplify the device currents. We have implemented a virtual lock-in amplifier using Labview (version 8.2) with the NI PCI 6259 M series DAQ card. This VLIA is just as good, if not better, than any of the off the shelf commercial lock-in amplifiers. All software is implemented in Labview and is run on a Dell Optiplex GX270 desktop computer with a 3.2 GHz processor and 1 GB of RAM.

Some of the noise sources that are associated with our EDMR measurements include quantization noise, the ambient noise from the surrounding hardware, and most importantly, the internal shot, thermal, and flicker noise arising from within the device under observation [6,7]. Fig. 1 illustrates the signal path in a typical EDMR spectrometer and shows the types of noise introduced to the signal as well as where it is added. The sampler represents the last contribution of noise because noise cannot be added to the signal once it is digitized. As mentioned earlier, we utilize a digital lock-in amplifier which removes most of the preamp and ambient noise but adds quantization noise (because it digitizes the signal before it is processed). The noise due to sources internal to the device are not entirely removed at the lock-in stage because this noise is to some extent present at the modulation frequency and lock-in phase. Therefore, it is the noise internal to the device that primarily determines the SNR at the output of the lock-in amplifier.

Fig. 2 illustrates current noise spectra from a MOSFETs configured as a gate controlled diode for three different biasing conditions. The top figure represents the condition with 0 V applied to the source and drain of the MOSFET; the current flowing through the device is thus essentially zero, therefore, this indicates that the noise spectrum observed in this case is the noise generated by the preamp. Note that this is more or less a white spectrum, meaning that the noise variance at all frequencies is the same. The middle figure represents the condition in which a small forward bias is applied to the source and drain of the MOSFET, yielding a dc current of 0.002 μA. The bottom figure illustrates a condition in which a large forward bias is applies to the source and drain yielding a dc current of 5 μA. This latter configuration corresponds to the biasing condition which results in maximum recombination and the operating point of our EDMR experiments. Note that this spectrum is significantly different than the other two. The reason for this is because of the significant flicker and shot noise that is introduced with larger dc currents [6]. This indicates that the dominating source of noise in the EDMR experiment is due to flicker and shot noise and that the noise from the preamp only becomes a problem when smaller devices (smaller currents) are being used.

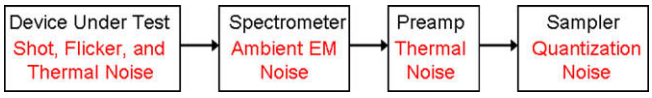

Fig. 1. Signal path in our EDMR spectrometer illustrating the types of introduced as well as the sources.

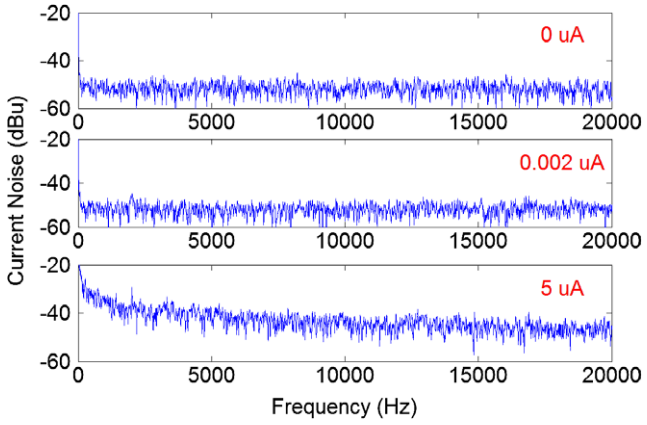

Fig. 2. Current noise spectra from a 4H SiC MOSFET configured in a gated controlled diode biased with three different voltages.

Initially, we attempted to reduce the noise observed in the EMDR experiments with adaptive noise cancellation techniques with a field programmable gate array (FPGA) before lock-in detection. The logic of processing EDMR signals before lock-in detection was the hope that a better representation (ie: improved SNR) of the amplitude modulated input signal would result in an improved SNR signal at the output of the LIA. It turned out that only minimal improvement was achieved because, as mentioned earlier, the majority of the noise in the EDMR measurement arises from the device under study and not the surrounding ambient noise. Also, lock-in detection itself is an extremely effective means of removing noise because it is not only frequency sensitive, but it is sensitive to phase as well. Therefore, the only noise that contaminates the post-lock-in EDMR signal is the noise with a frequency content near that of the modulation frequency. As a result, we decided to move our search to the output of the LIA for an effective way to enhance the sensitivity of EDMR.

In some cases, the devices under study have very few defects which make signal acquisition very difficult and time consuming. We have developed a way to expedite the averaging process by utilizing the predictability of the autoregressive noise features at the output of the LIA. (The time constant of the LIA determines the correlation between successive samples and hence, the predictability). We term this tool an adaptive signal averager (ASA) which utilizes adaptive linear prediction as illustrated in Fig. 2. It works by using the conventional scan average as the desired response in an adaptive linear prediction configuration. The linear predictor $\boldsymbol{w}_n$ is a finite impulse response (FIR) filter of length *p*and the input to the linear predictor is the tapped delayed noisy EDMR vector $\boldsymbol{x}(n)$ also of length *p*, where *n* is the present time index. These vectors are represented as column vectors which are indicated by the transpose operators *T*. The input samples $x(n)$ of the vector $\boldsymbol{x}(n)$ are composed of the desired EDMR signal $d(n)$ and an arbitrary noise component $u(n)$.

$$\boldsymbol{w}_n = [w_n(1), w_n(2), \ldots, w_n(p)]^{\mathrm{T}} \quad (1)$$

$$\boldsymbol{x}(n) = [x(n-1), x(n-2), \ldots, x(n-p)]^{\mathrm{T}} \quad (2)$$

The tapped delayed input vector $\boldsymbol{x}(n)$ is analogous to a shift register. First, the vector is initialized to the first *p* samples of the EDMR signal. Then, when a new sample is acquired, the samples are shifted to make room for the present sample. As a result, the oldest sample is forced out of the input array. This shifting process is then continued until the end of the scan. The present sample of the tapped delayed input vector is represented by the term $x(n-1)$ which is counterintuitive because this notation implies that it is the first past sample. This notation is used because we are attempting to predict the future sample $d(n)$ based on past values of the noisy input sam-

ples $x(n-1)$ to $x(n-p)$. This is why this technique is termed linear prediction. The prediction or estimate $d_{est}(n)$ of the desired signal is simply computed by the inner product of these two vectors.

$$d_{est}(n) = \boldsymbol{x}^T(n)\boldsymbol{w}_n \tag{3}$$

The estimate is then subtracted from the scan average to form an instantaneous error $e(n)$ which is used in an algorithm to update the weights of the FIR predictor.

$$e(n) = d(n) - d_{est}(n) \tag{4}$$

There are many forms of adaptive filters but the two most widely used and efficient are the least mean squares (LMS) and recursive least squares (RLS) adaptive filters. These filters are advantageous because they are capable of tracking non-stationary signals and noise and neither algorithm requires an estimate of the signal or noise statistics. This is desired for EDMR experiments because these statistics differ significantly from device to device and vary over time. The main advantage that the RLS algorithm has over the LMS algorithm is that it has about an order of magnitude faster convergence time, though, in most cases, the LMS algorithm is known to have better tracking performance [9]. To increase the tracking performance of the RLS algorithm, we utilized the exponentially weighted RLS (EWRLS) algorithm by incorporating an exponentially weighing factor λ into the system. By doing this, the algorithm effectively becomes more sensitive to changes in the noise environment. The exponential weighting factor λ controls the memory of the system and is chosen to be in the range $0 < \lambda < 1$. The EWRLS algorithm becomes the RLS algorithm when λ is chosen to be 1 which provides the system with infinite memory. This implementation is usually undesirable for EDMR measurements with moderate to high SNR because the filtering performance is degraded. This is because the correlation matrix cannot update fast enough when transitioning from noise to signal and vise versa. Therefore, the exponential weighting factor should be kept slightly less than one for signals with moderate to high SNR (see Fig. 3).

The EWRLS algorithm attempts to minimize the exponentially weighted sum of squared errors cost function which is given by Eq. (5).

$$\xi(n) = \sum_{i=0}^{n} \lambda^{n-i}|e(n)|^2 \tag{5}$$

In order to minimize this cost function, the gradient is taken with respect to the weights of the FIR predictor and set equal to zero which is given by Eq. (6).

$$\nabla\xi(n) = -\sum_{i=0}^{n} \lambda^{n-i}\boldsymbol{x}(i)e(i) = 0 \tag{6}$$

This resultant vector represents the direction of steepest decent on the sum of squared error surface. Plugging in for the error and rearranging yields the set of linear equations given in Eq. (7).

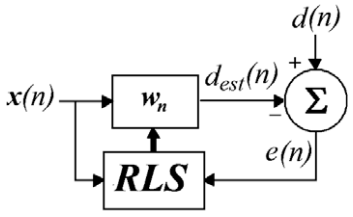

Fig. 3. Block diagram of the adaptive linear predictor using the EWRLS algorithm. $\boldsymbol{x}(n)$ represents the input vector composed of samples $x(n-1)$ to $x(n-p)$ and $\boldsymbol{w}_n$ represents the FIR predictor of length p. The predictor output $d_{est}(n)$ is subtracted from the desired signal $d(n)$ (the conventional average) to form an instantaneous error $e(n)$. This error, along with the input vector, is fed into the RLS algorithm to update the weights of the filter so that a better prediction can occur when the next sample is presented to the system.

$$\left[\sum_{i=0}^{n} \lambda^{n-i}\boldsymbol{x}(i)\boldsymbol{x}^T(i)\right]\boldsymbol{w}_n = \sum_{i=0}^{n} \lambda^{n-i}d(i)\boldsymbol{x}(i) \tag{7}$$

This result can be simplified by realizing that the terms in the brackets on the left is the summation of exponentially weighted deterministic autocorrelation matrices $\boldsymbol{R}_x(n)$ of the input signal from time index $0 < i < n$ and the right hand side is the summation of exponentially weighted deterministic cross correlation vectors $\boldsymbol{r}_{dx}(n)$ of the desired signal and the input signal from time index $0 < i < n$. With this realization, Eq. (7) in matrix form is equivalent to Eq. (8).

$$\boldsymbol{R}_x(n)\boldsymbol{w}_n = \boldsymbol{r}_{dx}(n) \tag{8}$$

Therefore, the weight vector $\boldsymbol{w}_n$ is found by multiplying the cross correlation vector $\boldsymbol{r}_{dx}(n)$ with the inverse correlation matrix $\boldsymbol{R}_x^{-1}(n)$. Calculation of this inverse is computationally intense so it is not desirable to calculate it every time a new sample is presented to the system. Therefore, one way to reduce the computational time is to realize that $\boldsymbol{R}_x(n)$ and $\boldsymbol{R}_x^{-1}(n)$ can be solved recursively. It can be easily shown that,

$$\boldsymbol{R}_x(n) = \lambda\boldsymbol{R}_x(n-1) + \boldsymbol{x}(n)x^T(n) \tag{9}$$

Now that $\boldsymbol{R}_x(n)$ can be solved for in terms of $\boldsymbol{R}_x(n-1)$, there needs to be a way to compute the inverse of this matrix. This is called the matrix inversion lemma. The inverse of the exponentially weighted autocorrelation matrix in Eq. (9) can be solved using Woodbury's identity [9,10]. Woodbury's identity states that matrix A of Eq. (10) can be inverted with the relation shown in Eq. (11). This identity only holds if A and B are positive-definite p-by-p matrices, D is a positive-definite n-by-p matrix, and C is an p-by-n matrix. The relation is easily shown by computing $AA^{-1} = I$, where I is the identity matrix.

$$A = B^{-1} + CD^{-1}C^T \tag{10}$$

$$A^{-1} = B - BC(D + C^TBC)^{-1}C^TB \tag{11}$$

Note that the following derivation is for real valued data. The transpose operations would be replaced with the hermitian operator for imaginary valued data. Comparing Eqs. (10) and (11), it can be realized that

$$A = \boldsymbol{R}_x(n) \tag{12}$$

$$B^{-1} = \lambda\boldsymbol{R}_x(n-1) \tag{13}$$

$$C = \boldsymbol{x}(n) \tag{14}$$

$$D = 1 \tag{15}$$

Then, plugging Eqs. (12)–(15) into (11), the exponentially weighted inverse autocorrelation matrix can be computed recursively as follows.

$$\boldsymbol{R}_x^{-1}(n) = \lambda^{-1}\boldsymbol{R}_x^{-1}(n-1) + \frac{\lambda^{-2}\boldsymbol{R}_x^{-1}(n-1)\boldsymbol{x}(n)\boldsymbol{x}^T(n)\boldsymbol{R}_x^{-1}(n-1)}{1 + \lambda^{-1}\boldsymbol{x}^T(n)\boldsymbol{R}_x^{-1}(n-1)\boldsymbol{x}(n)} \tag{16}$$

This equation is usually reduced into simpler form, as shown in Eq. (17)

$$\boldsymbol{R}_x^{-1}(n) = \frac{1}{\lambda}[\boldsymbol{R}_x^{-1}(n-1) - \boldsymbol{g}(n)\boldsymbol{z}^T(n)] \tag{17}$$

where,

$$\boldsymbol{z}(n) = \boldsymbol{R}_x^{-1}(n-1)\boldsymbol{x}(n) \tag{18}$$

$$\boldsymbol{g}(n) = \frac{1}{\lambda + \boldsymbol{x}^T(n)\boldsymbol{z}(n)}z(n) = \boldsymbol{R}_x^{-1}(n)\boldsymbol{x}(n) \tag{19}$$

The next step is to solve for the weight update. As stated earlier, the weight vector is found by multiplying the cross correlation vector $\boldsymbol{r}_{dx}(n)$ with the inverse correlation matrix $\boldsymbol{R}_x^{-1}(n)$. To reduce computation, $\boldsymbol{r}_{dx}(n)$ is solved recursively in a similar fashion to that of $\boldsymbol{R}_x(n)$ and is shown below.

$$\boldsymbol{r}_{dx}(n) = \lambda \boldsymbol{r}_{dx}(n-1) + d(n)\boldsymbol{x}(n) \tag{20}$$

The weight vector is found by computing the product of the autocorrelation matrix $\boldsymbol{R}_x^{-1}(n)$ obtained in Eq. (16) and the recursive cross correlation vector $\boldsymbol{r}_{dx}(n)$ formed by Eq. (21) and realizing that $\boldsymbol{R}_x^{-1}(n-1)\boldsymbol{w}_{n-1} = \boldsymbol{r}_{dx}(n-1)$.

$$\boldsymbol{w}_n = \boldsymbol{R}_x^{-1}(n)\boldsymbol{r}_{dx}(n) = \boldsymbol{w}_{n-1} + \boldsymbol{g}(n)\alpha(n) \tag{21}$$

where $\boldsymbol{g}(n)$ was defined previously and $\alpha(n)$ is the a priori error. The priori error is the error that occurs when using the previous set of filter coefficients $\boldsymbol{w}_n$ and is shown below,

$$\alpha(n) = d(n) - \boldsymbol{x}^T(n)\boldsymbol{w}_{n-1} \tag{22}$$

It is easy to see that the computation has been reduced significantly from the conventional LS algorithm because of the recursive nature of the autocorrelation and cross correlation functions. $\boldsymbol{R}_x^{-1}(n)$ can be initialized directly or by forming the matrix $\delta\cdot\mathbf{I}$, where δ is a constant called the regularization parameter and $\mathbf{I}$ is the identity matrix. The initialization of δ depends on the SNR of the signal under observation and should be calculated with the following [9,13].

$$\delta = \sigma_u^2(1-\lambda)^{\alpha} \tag{23}$$

where σ_u^2 represents the noise variance of an individual EDMR scan, λ is the exponential weighting factor, and α is a constant to be determined by the SNR of the EDMR scan. The parameter α should be chosen to be 1 for SNR > 30 dB, $-1 < \alpha < 0$ for SNR $\sim$ 10 dB, $\alpha < -1$ for SNR < -10 dB [9,13].

In some cases, the RLS algorithm can become unstable due to its mathematical formulation. This occurs when the inverse autocorrelation matrix loses its symmetry property [9,11]. This can be avoided simply by calculating the lower (or upper) triangle of the inverse autocorrelation matrix and filling the upper (or lower) triangle to preserve its symmetry property [9,12]. This technique is attractive not only because it prevents instability, but it also reduces computation. We utilized this method because we initially encountered instability problems.

As mentioned earlier, the ASA filters each incoming EDMR scan in real time via the EWRLS algorithm. The conventional average is used as the desired signal in the algorithm and can be thought of as an approximate guide for the filter to follow. Therefore, the filter allows the noise that it sees to pass, but it effectively reduces the variance of it thereby acting as a low pass filter with a time constant proportional to $(1-\lambda)^{-1}$ [9]. It is important to note that even though this adaptive filter may have a similar Fourier transform compared to that of a conventional low pass filter, it differs in many ways. A conventional low pass filter simply does not allow the input signal to pass any frequency greater than the cutoff frequency of the filter. The adaptive filter also acts as a frequency blocker but differs from a low pass filter because its weights are adapted so as to minimize the error between the desired signal and its prediction. As result, adaptive filters are capable of limiting the variance of the input signal relative to the desired signal. The filter only allows the variance of the input signal to deviate from the desired signal within a certain threshold. This is precisely why adaptive filters have superior performance to that of conventional low pass filters.

This filter is ideal for magnetic resonance experiments because one usually sacrifices a smaller time constant for the observation of noisier signals. As a result, each individual spectrum will contain more noise and will require the need for longer signal averaging to obtain a reasonable SNR. The filtered output scans are then averaged separately. The underlying idea for this action is that because the noise of the filtered scans is reduced, the noise in the filtered average will be reduced faster than that of the noise in the conventional average.

In conventional signal averaging, assuming the noise has a Gaussian distribution and is independent and identically distributed (iid) with variance σ_u^2, the averaged noise variance σ_{uN}^2 is reduced by a factor of the number of scans N in the average as given in Eq. (14).

$$\sigma_{uN}^2 = \frac{\sigma_u^2}{N} \tag{24}$$

The reduction in noise of the ASA can be determined by analyzing the error that is introduced into the algorithm. For an individual scan, the error introduced into the system by the filter is the combination of the averaged noise in the conventional average $u_N(n)$ with variance σ_{uN}^2 and the prediction error of the filter $v(n)$.

$$e(n) = d(n) - d_{\text{est}}(n) = [d(n) + u_N(n)] - [d(n) + v(n)] = u_N(n) - v(n) \tag{25}$$

For ease of analysis, it is assumed that the prediction error is also Gaussian random variable and has 0 mean and variance σ_v^2. Therefore, the variance of the error σ_e^2 for an individual scan is found by adding the variances of each of the random variables.

$$\sigma_e^2 = \sigma_{uN}^2 + \sigma_v^2 = \frac{\sigma_u^2}{N} + \sigma_v^2 \tag{26}$$

If M filtered scans are averaged, then the reduction in noise variance achieved by the ASA is simply given in Eq. (17).

$$\sigma_{e_M}^2 = \frac{\sigma_u^2}{NM} + \frac{{\sigma_v}^2}{M} \tag{27}$$

where $M < N$. The reason M scans are averaged and not N is because we want the conventional average to build up a reasonable desired response before the filter is applied so a better prediction can be achieved. N is not that much greater than M so they are approximately equal when considering longer averages. Therefore, as N and M get larger, the faster the first term in (17) dies away which implies that the dominating source of noise will eventually be due only to the prediction error of the filter. This is desirable because it is this first term that actually slightly biases the ASA. By allowing the conventional average to build a reasonable desired signal before the filter is applied, the noise bias is gradually removed. It turns out that not many scans are required to be averaged for this bias to be removed.

Another concern one might have would be whether or not the adaptive filter would have any affect on the g value. It turns out that this technique does not introduce any significant shift in the spectrum, which would be expressed in a shift in the g value. The reason for this is because an FIR filter of small length (<256 weights) is used in our implementation. FIR filters have a linear phase and introduce a small constant delay at the output of the filter. This delay is actually filter length dependent so that longer delays are encountered for filters of longer length. Though, this delay for longer length filters is not significant enough to introduce any significant error in the g value measurement. This was confirmed when a relatively long filter (1024 weights) was used in the algorithm to find the g value of a certain SiC MOSFET. The g values from the unconventional average differed from the filtered average by about 0.000001. In our experiments, we usually represent the g value with a much larger source of error of ±0.0003. Therefore, the small error introduced by the filter is negligible.

As discussed earlier, the prediction of the desired signal is always better than or equal to that of the noisy input because the filter is optimized to minimize the sum of squared errors. Therefore, the reduction in noise of the filtered average will always be better than that of the original average over time, despite being averaged with fewer scans. As a result, one can see why this averaging process is expedited; averaging a random variable with a small variance (prediction error) will converge much faster than averaging a random variable with larger variance (noise error).

The EWRLS ASA was implemented in Labview version 8.2 software and applied to EDMR for 4H SiC MOSFETs. The spectrometer

settings used in the scan were purposely chosen to reduce the SNR of the signal so as to better visually observe the improvement of the filtered signal. The variables that were used in the EWRLS algorithm were $\lambda = 0.98$, $\delta = 1$, $p = 32$ weights, and the filter was applied after averaging 15 scans. The reason 15 scans were averaged before the filter was applied was so that the filter had a better represented spectrum (moderate SNR as opposed to low SNR) upon which to base its prediction. Also, 15 scans in the conventional average were enough to remove the noise bias which was discussed previously.

Fig. 4 compares the performance of the filter of an individual scan. With the signal amplitude normalize to 1, the noise variance was calculated to be $\sigma_u^2 = 0.0315$ in the unfiltered trace and was calculated to be $\sigma_u^2 = 0.00278$ for the filtered trace. (These values were calculated by taking the variance of the difference between the individual scan and the final average.) As a result, an 11.3 times reduction in noise variance was observed in a single scan which corresponds to a 11.3 times reduction in time as well. Fig. 5 compares the average of 100 unfiltered scans and the average of 85 filtered scans. Note that the filtered average is not as noisy as the conventional average and has almost converged to its final value. Fig. 6 compares the average of 1000 unfiltered scans and the average of 985 filtered scans. Note that significant noise is present in the unfiltered average whereas the noise is not visually observable in the filtered average. Also, the variance of the noise that remains in the conventional average after 1000 scans is approximately equal to the noise variance in the filtered average after about 90 or so scans as illustrated in Fig. 7. As a result, the reduction in a

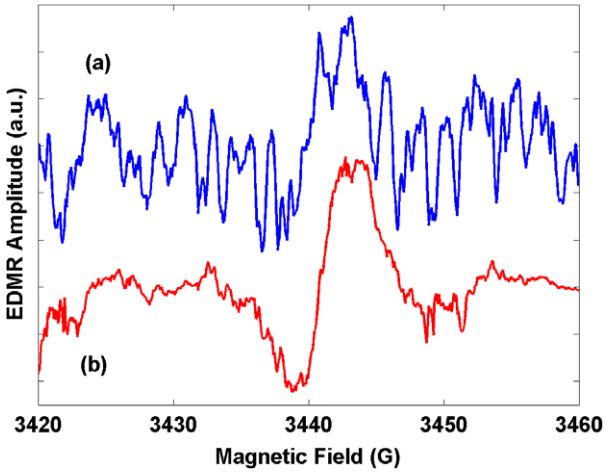

Fig. 4. Individual unfiltered scan (a) compared to the individual filtered scan (b).

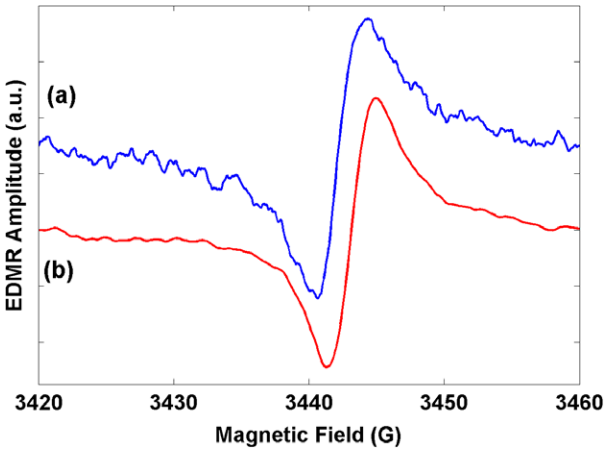

Fig. 5. Average of 100 unfiltered scans (a) compared to the average of 85 filtered scans (b).

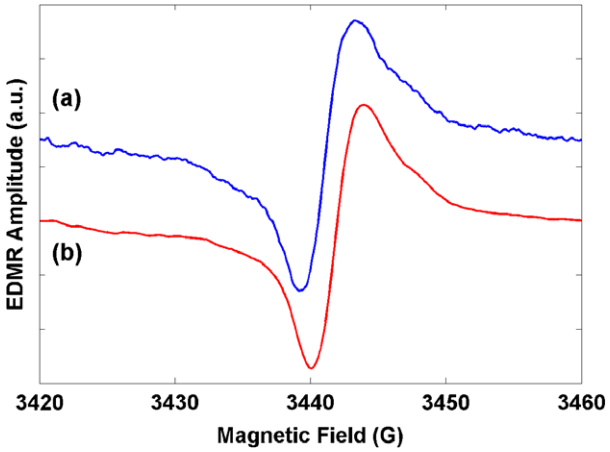

Fig. 6. Average of 1000 unfiltered scans (a) compared to the average of 985 filtered scans (b).

noise variance by factor of 11.3 in an individual scan is equivalent to a reduction in time by the same amount as illustrated in Fig. 6. In this particular experiment, the conventional average (1000 scans at 1 min each) took 1000 min to complete. The filtered average converged in approximately 90 scans which amounts to 910 less minutes of scanning time to obtain a comparable SNR. This is a very significant consideration, especially for measurements that require days of signal averaging. A signal that would usually require 10 days of signal averaging would be reduced to averaging for less than 1 day (assuming similar filter performance).

A concern one might have would be when to apply the filter. It turns out, that even if the SNR is less than 1, the filtered average will converge to the same result as the original average. This is ok to do so long as a sufficient number of scans are averaged first before the filter is applied to remove much of the noise bias, as discussed earlier. We applied the ASA to a 4H SiC bipolar junction transistor which has an EDMR spectrum which has weak sidepeaks due to hyperfine interactions with nearby ^{13}C and ^{29}Si. This hyperfine structure is unobservable until at least 20 or so scans in the average. The filter was applied after 15 scans (before any of the weak hyperfine structure was observed) and our results show that the unfiltered average and the filtered average are identical after 250 scans. The only difference is that the filtered average converged in many fewer scans.

The EWRLS ASA is an extremely useful and efficient tool for EDMR. It is capable of reducing the noise variance by a factor of 11.3 in a single trace and as expected, the average of the filtered scans was shown to converge by a similar factor. This filter is even

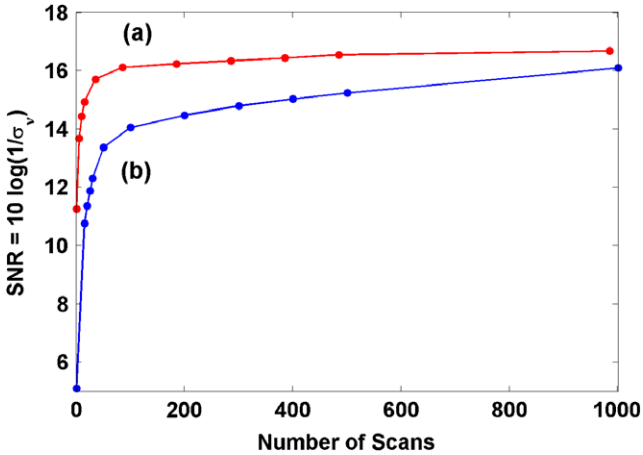

Fig. 7. SNR ratio the (a) filtered average and the (b) conventional average.

successful when the SNR of an EDMR scan is less than 1. With such great reduction in noise, the ASA effectively expedites the time of averaging. This tool not only has great potential in EDMR, but would also be beneficial for ESR in general. This approach may be very useful in quantum computing experiments where extensive signal averaging may be required for single spin detection. This filter can also be applied to any field where extremely high sensitive and relatively short acquisition times are required. Furthermore, it should be noted that this filter can also be applied to any field where signal averaging is utilized.

Acknowledgments

This work has been primarily supported by the US Army Research Laboratory, with additional support from Texas Instruments through Semiconductor Research Corporation Custom funding.

References

[1] D.J. Lepine, Phys. Rev. B 6 (1972) 436.
[2] D. Kaplan, I. Soloman, N.F. Mott, Civil Eng. Mag. 39 (1978) L51.
[3] J.P. Campbell, P.M. Lenahan, C.J. Cochrane, A.T. Krishnan, S. Krishnan, IEEE Trans. Dev. Mat. Rel. 7 (4) (2007) 540–557.
[4] C.J. Cochrane, P.M. Lenahan, A.J. Lelis, Appl. Phys. Lett. 90 (2007) 123501.
[5] J.A. Weil, J.R. Bolton, J.E. Wertz, Electron Paramagnetic Resonance: Elementary Theory and Practical Applications, John Wiley and Sons, Inc., New York, NY, 1994.
[6] Aldert van der Ziel, Noise in Solid State Devices and Circuits, Wiley, 1986.
[7] P.D. Petrova, TELSIKS 2 (2003) 577–580.
[8] S. Pajevic, G.H. Weiss, K.W. Fishbein, R.G.S. Spencer, Use of the Adaptive Line Enhancement Filter for SNR Improvement in NMR Spectroscopy, Proc. Intl. Sot. Mag. Reson. Med. 8 (2000) 1778.
[9] S. Haykin, Adaptive Filter Theory, fourth ed., Prentice Hall, Inc., 2002.
[10] M. Hayes, Statistical Digital Signal Processing and Modeling, John Wiley and Sons, Inc., 1996.
[11] M.H. Verhaegen, Automatica 25 (3) (1989) 437–444.
[12] B. Yang, IEEE Trans. Sig. Proc. 42 (12) (1994).
[13] G.V. Moustakides, IEEE Trans. Sig. Proc. 45 (10) (1997) 2468–2476.