**A TWO-DOMAIN REAL-TIME ALGORITHM FOR DATA REDUCTION
AND SNR OPTIMIZATION OF HIGH-RATE MEASUREMENTS
ON ACCELERATOR MAGNETS**

P. Arpaia^{1,2}, M. Buzio², V. Inglese^{2,3}

Abstract

A real-time algorithm of data reduction, based on the combination of two lossy techniques specifically optimized for high-rate magnetic measurements in two domains (e.g. time and space), is proposed. The first technique exploits an adaptive sampling rule based on the power estimation of the flux increments in order to optimize the information to be gathered for magnetic field analysis in real time. The tracking condition is defined by the target noise level in the Nyquist band required by post-processing procedure of magnetic analysis. The second technique uses a data reduction algorithm in order to improve the compression ratio while preserving the consistency of the measured signal. The allowed loss is set equal to the random noise level in the signal in order to force the loss and the noise to cancel rather than to add, by improving the signal-to-noise ratio. Numerical analysis and experimental results of on-field performance characterization and validation for two case studies of magnetic measurement systems for testing superconducting and resistive magnets of the Large Hadron Collider at the European Organization for Nuclear Research (CERN) are reported.

CERN-TE-2009-001
26/06/2009

1 Department of Engineering, University of Sannio, Benvento, Italy

2 CERN, Technology Department, Geneva, Switzerland

3 Department of Electrical Engineering, University of Naples, Napoli, Italy

To be published in "Measurement Science and Technology"

1. Introduction

At the European Organization for Nuclear Research (CERN), the tests on the about 10,000 magnets of the particle accelerator Large Hadron Collider (LHC) disclosed new horizons for the related magnetic measurements [1]. A new generation of fast transducers [2]-[3], capable of increasing by three orders of magnitude the bandwidth of harmonic measurements (10 to 100 Hz), when compared to the standard rotating coil technique (typically 0.1 Hz), and still maintaining a typical resolution of 10 ppm, have been developed. Moreover, also a Fast Digital Integrator (FDI), reducing flux acquisition time down to 4.00 μ s, within a resolution of 50 ns, was prototyped and field-tested successfully [4].

The higher sampling rate of such a measurement setup increases the amount of resulting data by producing an exponential rise in data storage requirements (typically about three orders of magnitude). Thus, data reduction algorithms are needed in order to decrease the size of measurement results by controlling the quality loss simultaneously.

Lossless compression algorithms, such as Huffman coding [5], arithmetic coding [6], or lossless JPEG [7], usually exploiting statistical redundancy for a more concise representation without error, are the first logical choice. However, when the original data contains sufficient redundancy, and a suitable assumption can be made on whether a certain approximation is not critical, better compression ratios are achieved by lossy methods [5]-[26]. Various approaches to dimensionality reduction such as principal components analysis, entropy measures for ranking features, and methods to discretize data were reviewed in [8]. In creating CAD geometry data from existing parts by surface laser scanning, uniform and non-uniform grid methods were proposed [9]. In biomedical applications, adaptive sampling algorithms are widely used to reduce data size at their source, while preserving clinical acceptability of the reconstructed signal [10]. In particular, reduction techniques belonging to the class of significant-point-extraction algorithms, such as Turning Point (TP) [11], Amplitude Zone Time Epoch Coding (AZTEC) [12], and the Fan [13], are proposed. These are lossy algorithms retaining only samples with important information about the signal. In machine learning and pattern recognition with high-dimensional data, unsupervised and supervised greedy reduction algorithms based on computing discernibility power of a crisp/fuzzy equivalence relation were proposed [14]. In knowledge discovery in databases, evolutionary instance selection algorithms were shown to consistently outperform non-evolutionary ones by better instance reduction rates, higher classification accuracy, and easier models [15]. In sensor networks for environmental monitoring, the distributed structure of the measurement system led to the development of decentralized autonomous cooperative intelligent sensors, capable of minimizing resource consumption (energy and network bandwidth), by adapting their sampling rates to the actual environmental conditions [16]-[19].

particle accelerator Large Hadron Collider (LHC) disclosed new horizons for the related magnetic measurements [1]. A new generation of fast transducers [2]-[3], capable of increasing by three orders of magnitude the bandwidth of harmonic measurements (10 to 100 Hz), when compared to the standard rotating coil technique (typically 0.1 Hz), and still maintaining a typical resolution of 10 ppm, have been developed. Moreover, also a Fast Digital Integrator (FDI), reducing flux acquisition time down to 4.00 μ s, within a resolution of 50 ns, was prototyped and field-tested successfully [4].

The higher sampling rate of such a measurement setup increases the amount of resulting data by producing an exponential rise in data storage requirements (typically about three orders of magnitude). Thus, data reduction algorithms are needed in order to decrease the size of measurement results by controlling the quality loss simultaneously.

Lossless compression algorithms, such as Huffman coding [5], arithmetic coding [6], or lossless JPEG [7], usually exploiting statistical redundancy for a more concise representation without error, are the first logical choice. However, when the original data contains sufficient redundancy, and a suitable assumption can be made on whether a certain approximation is not critical, better compression ratios are achieved by lossy methods [5]-[26]. Various approaches to dimensionality reduction such as principal components analysis, entropy measures for ranking features, and methods to discretize data were reviewed in [8]. In creating CAD geometry data from existing parts by surface laser scanning, uniform and non-uniform grid methods were proposed [9]. In biomedical applications, adaptive sampling algorithms are widely used to reduce data size at their source, while preserving clinical acceptability of the reconstructed signal [10]. In particular, reduction techniques belonging to the class of significant-point-extraction algorithms, such as Turning Point (TP) [11], Amplitude Zone Time Epoch Coding (AZTEC) [12], and the Fan [13], are proposed. These are lossy algorithms retaining only samples with important information about the signal. In machine learning and pattern recognition with high-dimensional data, unsupervised and supervised greedy reduction algorithms based on computing discernibility power of a crisp/fuzzy equivalence relation were proposed [14]. In knowledge discovery in databases, evolutionary instance selection algorithms were shown to consistently outperform non-evolutionary ones by better instance reduction rates, higher classification accuracy, and easier models [15]. In sensor networks for environmental monitoring, the distributed structure of the measurement system led to the development of decentralized autonomous cooperative intelligent sensors, capable of minimizing resource consumption (energy and network bandwidth), by adapting their sampling rates to the actual environmental conditions [16]-[19].

Anyway, in all these fields, applications have been developed so far mainly for measurements with low sampling rates (typically below 100 S/s). Furthermore, some of the proposed methods turn out to be highly complex, by requiring: a model of the measurand physical process [16], the design of a compound Kalman filter [17], or the solution of a constrained optimization problem [18]. Their computational burden limits their application to fast real-time data reduction above a few tens of samples per second. In [19], an *fft*-based algorithm for estimating the optimal sampling rate of snow capacitance sensor network is proposed. It is capable of reducing the sensor power consumption by tracking the dynamics of the process under monitoring and adapting the sampling frequency through the CUSUM [20] test in real-time. However, it has been used only for very low sampling rates (typically one sample every few tens of seconds), thus proving its effectiveness only for processes with a slow time variation. Other techniques, such as histogram equalization [21] and entropy-based adaptive sampling [22], were proposed for image processing applications. However, these methods include a phase of data analysis needing for the availability of the data set as a whole, thus preventing their use for fast real-time applications.

In this paper, a real-time two-domains data reduction algorithm based on adaptive sampling and significant point extraction, specifically optimized for high-rate magnetic measurements, is presented. In particular, in Section II, the proposed algorithm is illustrated by highlighting the basic ideas, the procedure, and main design criteria for magnetic flux measurements. In Section III, the numerical analysis of its performance is presented. Finally, in Section IV, experimental results of field characterization and validation obtained at CERN on two test benches for automatic measurements of superconducting and resistive magnets, respectively, are shown.

2. Proposal

In the following, (i) the *basic ideas*, (ii) the *strategy*, (iii) the *procedure*, and (iv) the *design criteria* of the proposed algorithm are illustrated.

2.1. Basic Ideas

The algorithm was based on the following design concepts:

1. Data are reduced by analyzing the results in real time during the measurement and by storing consequently only the necessary ones. Two different lossy techniques, applied in different domains (e.g. time and space) and suitable for the signal features in those domains, are applied.

The first technique, the *adaptive tracking sampler*, analyzes the acquired data on the basis of a Nyquist power threshold mechanism of tracking, in order to adapt the sampling rate for capturing only the significant features of the signal (Fig. 1). The second technique, the *noise-cancelling compressor*, is based on a significant point extraction: only the samples containing significant information are retained by specifying the maximum error allowed on the reduced signal.

2. In the *adaptive tracking sampler*, according to the fundamental data reduction principle of gathering only the minimum amount of information sufficient to any kind of further analysis, a rule for adapting the sampling by tracking the signal power in the Nyquist bandwidth is defined [19]. Under the assumption of signal stationarity with respect to the algorithm adaptation time, sampling rate is modified by analyzing the signal in the frequency domain through a general technique for determining its spectral content. The sampling rate is adapted in real time when a limit condition is approached. The limit condition is based on the sampling rate defining the Nyquist bandwidth including the minimum power necessary to capture the required features of the signal. The power is estimated in a frequency band whose upper limit is the Nyquist frequency, compared with thresholds representing the power level in the band for which the current sampling rate is considered adequate. In particular, the sampling rate is increased/decreased when relatively high-frequency terms are present/absent into the power estimator.
3. In the second technique, the *noise-cancelling compressor*, between two observations where the signal is monitored and the sampling rate adapted at evenly spaced time intervals, the resulting signal is decomposed in a set of linear problems, to which classic lossy algorithms for significant point extraction, such as the Fan [11], [13], are applied. A suitable mechanism to control the related error allows a proper trade-off among compression speed, compressed data size, and quality loss. In particular, under the assumption of random noise (i.e. white but not necessarily Gaussian) and deterministic signal, if the allowed loss is set equal to the noise level, loss and noise tend to cancel rather than to add, by increasing signal-to-noise ratio [24]. In other words, by means of a suitable design strategy, the loss mainly involves the noise content of the signal, by improving both the signal quality and the compression ratio simultaneously. This approach does not require a priori knowledge of the spectral properties of the noise-free signal.

2.2. Strategy

In the *adaptive tracking sampler*, if the sampling rate is f_s , the power estimator is unable to adequately

treat a signal with frequency approaching $f_s/2$. Moreover, if the sampling rate is only f_s , the presence of signals with frequency greater than $f_s/2$ can not be revealed. However, an increase of the estimated power in the observed frequency band $[f_0, f_s/2]$, with $f_0 < f_s/2$ to be suitably chosen, at a level above a suitable threshold, can reveal higher frequency components approaching the limit of the current observable band. Hence, an increase in resolvable frequencies can be obtained by triggering an increase in the sampling rate. Likewise, the absence of large terms in the band $[f_0, f_s/2]$ is used to trigger a decrease in the sampling rate. In this way, error penalties incurred by slightly reducing the sampling rate in places where the signal is relatively uninteresting are not likely to be high. The tracking mechanism is explained graphically in Fig. 2.

This is true if the variation of the power content of the signal can be considered stationary with respect to the algorithm adaptation time. The frequency f_0 is to be chosen as a function of f_s , thus both the extremes of the observed frequency band can vary while the sampling rate is adapted.

While the tracking mechanism adapts the sampling frequency by analyzing subsequent batches of data, the *noise-cancelling compressor* discards possible redundant points between each pair of consecutive observations. The information loss is controlled by means of the tolerance ε : the approximation maximum error. The *noise-cancelling compressor* operates on the acquired samples, by approximating them by means of straight line segments and removing redundant points along the way. Good results are obtained if the algorithm is applied to each of the linear data subsets. The end points of the segments are determined on the basis of a maximum-error criterion requiring that the results of an approximation always fall within a user-specified range. The error range can be adaptive and not necessarily symmetrical.

2.3 Procedure

The procedure of the proposed algorithm acts in two different domains by exploiting the *adaptive tracking sampler* and the *noise-cancelling compressor*, respectively.

The *adaptive sampling* algorithm can be summarized in the following steps (Fig. 3):

1. acquire a batch of samples at the current rate f_s ;
2. estimate the signal power in the frequency band $[f_0, f_s/2]$;
3. compare the estimated power to the threshold terms, and if necessary update f_s ;
4. if new data income then go to point 1, else exit.

The algorithm is modular and generic: different techniques can be used for power estimation, thresholds

computation, low-pass filtering, and sampling rate updating.

Subsequently, the *noise-cancelling compressor* is applied to each subset of linear data. Once the tolerance ε has been specified, the approximation method works on a discrete signal $f(k)$ through the following main steps [23]:

1. The starting point k is a non-redundant point by definition. Points k and $k+1$ are used to draw two straight lines starting at k and crossing $f(k+1)+\varepsilon$ and $f(k+1)-\varepsilon$, respectively, in order to check the redundancy of the point $k+1$.
2. The resulting lines are extended to point $k+2$. If point $k+2$ lies outside them, point $k+1$ becomes a non-redundant point and the process is repeated by starting from point $k+1$. If point $k+2$ lies between the lines, it is found to be redundant and point k is retained as the starting point.
3. Then, new lines are created by using points k and $k+2$. These lines are compared with the previous ones and the more restrictive are kept. If point $k+3$ lies outside the new lines, point $k+2$ becomes a non-redundant point and the process starts over with point $k+2$ as the starting point. If point $k+3$ lies between them, point $k+2$ also becomes redundant and new and narrower cones are created until a non-redundant point is found.

In this way, the original sampled signal is approximated by straight line segments sequentially. All the non-redundant points and all the lengths between each pair of consecutive non-redundant points are stored.

2.4. Design Criteria

In a magnet test application, the algorithm is to be designed by determining: (i) the stage where data reduction is applied, (ii) the technique for power estimation, (iii) the thresholds, and (iv) sampling updating.

Stage where data reduction is applied. In an automatic test system, data reduction techniques can be applied at the level of *instrument firmware*, *instrument software*, and *host PC software*. At the level of the *instrument firmware*, reduction acts on the device directly, thus reducing the data amount on the interface bus. Furthermore, an adaptive parametric management of the measurement process in real-time exploiting the computing capabilities of intelligent devices makes the instrument autonomous. However, access to the internal instrument code and programming is needed. At level of the *instrument software*, the reduction is achieved by a centralized computing unit. The adaptive management implies a longer adaptation time, therefore the risk of losing some information when sudden or unpredictable variations

arise is to be faced. At the level of *host PC software*, generally the sampling rate is not adapted. Data acquired at high sampling rate are temporarily stored in a buffer, analyzed and reduced before being definitively saved. The risk of losing some information is strongly reduced, but no optimization of the amount of data transferred through the bus is achieved.

For the present magnetic measurement application, application at the firmware level will be the long term goal. This work is concerned with application at instrument level, which allows much greater flexibility in changing algorithms and parameters and comparing results.

Technique for power estimation. The most straightforward technique for computing the power of the flux signal is the Fourier analysis. The proposed adaptive sampling updates the estimation of the signal power after a sample batch of fixed size is acquired.

Sampling rate updating. A suitable width of the frequency range considered for the power estimation is chosen equal to one forth of the resolvable bandwidth. Thus, when a power increase in the sampled flux is detected, the proposed adaptive sampling is likely to have enough time to increase the sampling rate before important signal features exit the current observable bandwidth. A suitable trade-off between reaction time of the algorithm to flux variations and computational burden was joined according to the specific application constraints.

Thresholds computation. The threshold terms are determined on the basis of the estimated noise level. In particular, the tolerance ϵ of the *noise-cancelling compressor*, and upper and lower limits on the power admissible in the observed frequency range are defined for the tracking mechanism of the *adaptive tracking sampler*, in order to trigger the increase/decrease of the sampling rate, respectively.

3. Numerical results

The algorithm was first tested in simulation, where ideality assumptions can be made on the input signal, by evaluating the effect of the frequency adaptation without other perturbation sources.

In the following, numerical results of (i) the *static tests*, and (ii) the *dynamic tests* of the algorithm are illustrated.

3.1 Static tests

The proposed algorithm was compared to the following state-of-the-art solutions [5]-[13]:

1. The TP, a straightforward algorithm producing a fixed reduction ratio of 2:1. It processes three

points at a time, stores the first one and retains one of the next two samples depending on which one preserves the turning point (i.e., the slope change) of the original signal.

2. The AZTEC algorithm decomposes raw sample points into plateaus and slopes, thus representing the original signal through a piecewise-linear approximation formed by a sequence of line segments.
3. The pure Fan algorithm, as said before, also uses a piecewise-linear approximation of the original signal, but unlike AZTEC draws lines between pairs of starting and ending points so that all intermediate samples are within some specified error tolerance.

Simulations were carried out on an ideal input signal defined on the basis of the actual signals acquired on the field in a typical setting of a rotating coil measurement of a superconducting dipole magnet. The signal was used as input to the simulator. After the reduction, some performance indices were estimated and subsequently compared to those computed on the original signal.

In static tests, the algorithm is fed with a constant frequency input signal, the sampling frequency is set to the optimal value, and then the reduced data acquired in steady state are used to estimate the indices of the reduced signal and to compare them with those of the input signal. Test results are summarized in Tab. I for different settings of the algorithms' parameters. The Compression Ratio (CR), defined as ratio between the size of original data and the size of compressed data, assesses the data reduction, while Signal-to-Noise And Distortion ratio (SINAD), Signal-to-Non Harmonic Ratio (SNHR), and Total Harmonic Distortion (THD) [27] the loss. Finally, the main component amplitude is evaluated to express the closeness of the reduced signal to the original one.

Results show that the *adaptive tracking sampler* provides the best tradeoff between CR and error on the reconstructed signal. As said before, a lossy data compression algorithm, when applied to a noisy signal with the allowed loss set equal to the noise strength, produces a filtered signal with reduced noise content [24]. In addition to that, since the *adaptive tracking sampler* operates a decimation with an Over Sampling Ratio (OSR) equal to the decimation ratio, the SNR is improved while reducing the data size.

The additional comparison of the *adaptive tracking sampler* with the pure Fan algorithm applied to sine wave signals typical of a rotating coils-based measurement system [28] highlighted satisfying results. Nevertheless, the Fan algorithm performs better than TP and AZTEC, in terms of compression capabilities and fidelity to the original signal, respectively.

Furthermore, the choice of the Fan algorithm inside the *noise-canceller compressor* was checked. By

exploiting the signal periodicity, the reduced sine wave can be decomposed in a set of quasi-linear time series, formed by points corresponding to the same angular position (Fig. 4). The results (Tab. II) highlight the remarkable tradeoff between compression ratio and approximation error of the Fan algorithm when applied to a data set representing a linear signal. Fig. 5 shows the corresponding rate-distortion plot [24], allowing the noise strength to be determined by means of a heuristic method. This is used for tuning the tolerance ε inside the *noise-canceller compressor*. In the conditions of Fig. 5, where a white noise of power σ^2 is applied, the method suggests the setting of ε to a value of 3σ .

3.2 Dynamic tests

Inside the *adaptive tracking sampler*, the proposed algorithm includes an *fft*-based tracking mechanism, working properly only if its adaptation time is negligible in comparison to the typical time variation of the signal. The algorithm tracking capabilities were verified for different settings of its parameters by a simulation with variable frequency input signal. Extremely fast changes were imposed by using step functions to define the time variation law of the input frequency. The algorithm had to follow this variation by updating suitably the sampling frequency. The algorithm dynamic performance is depicted in Fig. 6 for different values of α , the maximum change in the sampling frequency allowed in one step, expressed as a fraction of the current sampling frequency. The figure shows how adaptation times suitable for the typical magnetic measurements carried out at CERN can be achieved.

3.3 Algorithm performance

With the aim of checking the computational burden of the proposed solution, the algorithm performance were evaluated in simulation and compared to those of other state-of-the-art algorithms, by measuring the time for processing the same data set. Execution times were computed in MATLAB, with respect to a very easy and fast solution such as the AZTEC. Tab. III shows the result of the comparison, proving that the computation required suits the capabilities of a system typically used for the application to magnetic measurements.

4. Experimental results

The experimental proof of the principle and the performance assessment of the proposed algorithm were carried out at the magnet test facility SM18 of CERN by exploiting two demonstrators. Tests were carried out on dipole magnets with the rotating coil method (Fig. 7) [28], one of most accurate techniques for

magnet testing. In such a technique, a coil turns in the magnet under test and its output signal is proportional to the flux derivative, according to Faraday's law. The coil signal is acquired and integrated by a digital board in the angular domain thanks to the output pulses of an encoder mounted on the rotating coil shaft. Typically, the undesired harmonics in an accelerator magnet are at the level of $\leq 10^{-4}$ of the fundamental field [29], thus, particular attention has to be paid to their estimation. In practice, one coil is used to measure the main field, while a combination of coils in a so-called compensating configuration (i.e. cancelling out the main harmonic) is used to detect higher harmonics. Two flux signals are therefore produced, the absolute (main component) and the compensated (mostly other terms). An off-line Fourier analysis of these signals (CERN standard analysis [30]) finally yields the multipoles of the magnetic field generated by the magnet under test. The aim to be achieved is to reduce the data size and simultaneously to be able of reconstructing the field with an acceptable approximation. In the most demanding applications, this means estimating the main field harmonic, in module and phase, with an error lower than 10^{-4} - 10^{-5} T, and the higher harmonics with an error of few ppm of the main component.

Methods for data reduction through transformation were proposed [31]: data are compressed through a compact representation in a transformed domain, with a simultaneous noise suppression. The proposed algorithm exploits an *fft* transformation, and could be easily adapted to include such a mechanism for the rotating coils application. However, data reduction would be achieved only by releasing the flux harmonics at each fixed number of time samples, typically at every coil turn. Therefore, the instantaneous behaviour of the original signal can not be reconstructed and the resulting data can not be used for an even slightly different analysis. For this reason, a more general solution based on reducing and storing raw data in time domain was preferred.

Rotating coil measurements have some peculiar features useful for customizing the proposed algorithm implementation and improve its performance. In particular, the main field component and its first m harmonics have to be found, thus inside the *adaptive tracking sampler*, the tracking is tuned to detect the main component frequency f_1 , and the sampling rate is updated in order to extend the observable spectrum up to $m f_1$. Moreover, field harmonics have a spatial period along the circumference described by the coils during the measurement. The harmonics time period is therefore a function of their spatial period and of the motor rotation speed. Formally, for the m^{th} harmonic:

$$f_m = m \omega / (2\pi) \quad (1)$$

where ω is the angular speed in rad/s. On the other hand, as explained before, the flux sampling frequency is the trigger frequency, obtained by multiplying the number of points per turn of the angular trigger by the motor rotation speed:

$$f_s = N \omega / (2\pi) \quad (2)$$

where N is the number of flux samples acquired per turn. In other words, both harmonic frequencies and sampling frequency scale according to the same proportionality law with the coils shaft angular speed. Their ratio is therefore fixed by the number of samples per turn for any motor speed. In this case, according to the Nyquist criterion, the minimum number of points per turn required for the m^{th} harmonic to be resolved can be determined easily:

$$f_s \geq c m f_1, \text{ with } c \geq 2 \quad (3)$$

It is therefore possible to check the algorithm capability of approaching the theoretical optimal value.

Consequently, the *adaptive tracking sampler* operates in the angular domain, basically by adapting the number of points per turn to the signal characteristics (Fig. 8). The signal under analysis represents the flux over subsequent angular sectors, with width depending on the flux sampling frequency. Therefore, the sampling frequency adaptation implies a variation of the angular sectors where the flux increments are computed. Each of these values is obtained by integrating the voltage samples coming from a fixed rate ADC, thus the adaptation of the flux sampling frequency results in different OSR of the flux signal [4].

Then, the signal features allow the *noise-cancelling compressor* to be applied. The rotating coils-based technique typically produces a sinusoidal output. As shown in Fig. 8, if the series $\Delta\varphi_n(\theta_i)$ are considered as formed by flux samples corresponding to a given angular position (θ_i) and acquired for increasing time (expressed as the current turn number n), when the magnet is supplied with a constant or ramped current, they have a nearly linear pattern since they represent the flux variation on the same angular sector in different time instants. For each point along the circle, there will be an almost linear time series $\Delta\varphi_n(\theta_i)$ to which the *noise-cancelling compressor* can be applied with satisfying results.

The combined reduction procedure as a whole can be summarized:

1. the *adaptive tracking sampler* is let run in order to reduce to its optimal value, compatibly with the constraints, the number of samples acquired on every single turn in correspondence to the angular positions θ_i ;
2. for each one of these θ_i , a quasi-linear series $\Delta\varphi_n(\theta_i)$ is obtained, containing a point for each coil turn, and the *noise-cancelling compressor* is applied.

The signal is therefore reduced twice, in angular and time domain, thus increasing the overall achievable compression. This combined approach is capable of attaining a remarkable tradeoff between reduction ratio and fidelity of the reconstructed signal to the original one, since its two steps operate subsequent compressions by means of mechanisms fitting the signal features in their respective application domains. Finally, from a computational point of view it has to be noted that, since every series $\Delta\varphi_n(\theta_l)$ has one point for each coil turn, at the maximum allowed rotation speed (8 rps) each of them requires to process only 8 S/s, thus keeping the additional computing power required to the measurement system within reasonable limits.

Twofold campaigns of on-field tests were carried out on the proposed algorithm mainly aimed at (i) the *performance characterization*, and (ii) the *data reduction validation*.

4.1 Performance Characterization

A test bench aimed at validating on field the proposed algorithm was assembled in the CERN test facility SM18, by means of:

1. a motor controller MAXON EPOS 24, accessible through RS232, for handling the motor turning the coil inside a superconducting magnet at a constant rotation rate;
2. a Fast Digital Integrator (FDI), a CERN-University of Sannio proprietary PXI general-purpose digitalization board, configured for the coil signal acquisition and numerical integration [4];
3. an encoder board: a CERN proprietary PXI board, for managing the encoder pulses and feeding the trigger input of the FDI;
4. a superconducting magnet at cryogenic temperature (1.9 K) used as *unit under test*, supplied with a current of 1500 A to generate the magnetic field;
5. a software developed by means of the CERN-University of Sannio proprietary Flexible Framework for Magnetic Measurements (FFMM) [32].

The proposed algorithm was characterized on the field for (i) *performance analysis*, by varying its settings, and (ii) *improvement assessment*, in comparison to state-of-the-art algorithms.

As far as the *performance analysis* is concerned, the *adaptive tracking sampler* was executed off line in order to find the optimal sampling frequency rate for carrying out the flux measurement, and the signal was consequently reduced and saved. Afterwards, the multipole expansion of the magnetic field was determined by means of the CERN standard analysis process [16]. The procedure was repeated for

different settings of the algorithm, corresponding to different optimal sampling rates and consequently to different compression ratios.

As an example, the results with the Fast Digital Integrator acquiring 128 points per turn from a coil rotating at a speed of 8 rps are shown in Tab. IV, in reference to the original signal. Amplitude and phase of the main component of the magnetic field were considered, along with harmonics of the multipole expansion [29] up to order 10. The latter are normalized with respect to the main field component and multiplied by a factor 10^6 (i.e. expressed in ppm). These results highlight how the algorithm is capable of achieving a remarkable trade off between reduction ratio and fidelity of the reconstructed signal.

As far as the *improvement assessment* is concerned, the reduction of the same sinusoidal measurement data by means of the *adaptive tracking sampler* and a classic reduction method, the Fan algorithm [23], were compared. In Tab. IV, the RMS error on the multipole expansion obtained from both the algorithms, are reported for different compression ratios. For similar ratios, the comparison shows a remarkable reduction of the error when the *adaptive tracking sampler* is employed. In particular, with a compression ratio of 4 the proposed adapting sampling algorithm is still able to provide an estimation of the field harmonics within an RMS error of a few tenths of ppm.

The *noise-cancelling compressor*, according to combined approach explained before, was applied to the series $\Delta\varphi_n(\theta_l)$ obtained after reducing the data by means of the *adaptive tracking sampler*. The results achieved after the second step of reduction are reported in Tab. V, in terms of compression ratio and RMS error on the computed harmonics. The results highlight, for similar compression ratio, the significant performance improvement of the *noise-cancelling compressor* if applied to a linear signal with respect to the reduction of a sine wave reported in Tab. IV. The value of the tolerance ε was chosen on the basis of the noise level of the flux signals. Obviously different values have to be set for the absolute and compensated signal. In particular, the noise level estimated on the compensated signal is two orders of magnitude lower than on the absolute signal ($\sim 10^{-7}$ and $\sim 10^{-5}$ Vs, respectively), according to typical values observed on these signals.

4.2. Data reduction validation

The data reduction algorithm was subsequently validated on the same calibration bench with a reference resistive dipole TE1 magnet at room temperature, used as *unit under test*, supplied with a current up to 200 A (Fig. 9).

The real-time implementation of the algorithm was realized in C++, at driver level, inside the FFMM [32]. A software component, the class *DataReducer*, with its methods and attributes, was added to the framework. The data reduction can therefore be used through the interface of such module, without taking care of its actual implementation. In this way it is possible to change implementation or even include other compression algorithm, minimizing the impact of the changes required in the other modules of the framework FFMM. For the application based on rotating coils, the user just has to set an option in the script to activate the reduction of the data acquired by the digital integrators. The threads taking care of the data transfer from the FDI buffer to the disk were modified, so that the flow of data coming from the FDIs is intercepted and reduced by *DataReducer* before being saved on the disk (Fig. 10). The reduced data are also analyzed in order to update the compression ratio according to the current signal features.

First, the reduction was performed during an acquisition with the magnet supplied by a constant current, with the aim (i) of determining the noise levels of the absolute and compensated signals in order to tune the tolerance ε , and (ii) of testing the algorithm in stationary conditions. The results are shown in Tab. VI. The values of the parameter ε were chosen according to the noise level of the flux signal. The magnet was supplied with a current of 200 A, 4096 points per turn were initially set. The tracking mechanism, used to detect harmonics up to the 15th, reduced the number of points per turn to 256, thus achieving a compression ratio of 16. This value is in accordance with the theoretical value derived by considering eq. (3) with $m = 15$ and $c = 10$. In fact, this yields

$$\text{Number of points per turn} = f_s / f_1 \geq c n = 150 \quad (4)$$

Since the number of points per turn is to be a power of 2, its minimum value is 256.

Subsequently, the *noise-canceller compressor* further reduces the obtained data with a compression ratio of 3.4 and 3.7 for the absolute and compensated signals, respectively. On the whole, the proposed technique proves therefore to be capable of achieving the remarkable compression ratios of 54.4 and 59.2 for the absolute and compensated signal, introducing an acceptable error on the main component and on the harmonics of the resulting signals.

Afterwards, the compression was performed on the data acquired during an acquisition with the current in the magnet ramping at a rate of 10 A/s. The Fan algorithm was executed with the values of ε determined before. Tab. VII shows the results obtained in these measurement conditions. The figures in table highlight the good capabilities of the proposed approach also in non-stationary conditions. In particular, in this case a higher compression ratio is obtained with a smaller approximation error. The reason for that

could be that, when the current is low at the beginning of the ramp, the signal is significantly affected by the noise. The results therefore highlights that the algorithm is capable of filtering most of the noise, thus achieving high compression without increasing the related error.

5. Conclusions

A batch algorithm for data reduction through adaptive sampling and significant point extraction has been proposed. It is fast, reliable, cost effective and can be implemented in real-time in order to extract significant points from a sampled signal, with an information loss within an error range specified by the user through a threshold mechanism. The algorithm is modular and generic: different techniques can be used for power estimation, thresholds computation, low-pass filtering, and sampling rate updating.

Preliminary experimental results on magnetic measurements performed at CERN through the rotating coils technique proved the effectiveness of the algorithm. These results were further investigated by implementing the algorithm inside the measurement station in order to verify actual real-time performance. In addition, a comprehensive analysis was carried out also for the estimation of the magnetic field multipole expansion through rotating coils measurements. The tests were performed at the CERN test facility SM18 both on a superconducting LHC dipole and on a calibration resistive dipole magnet. The results highlight the remarkable tradeoff between fidelity to the original signal and achieved data reduction allowed by the proposed two-domain combined approach.

Previous results highlighted a synergic interaction between the two steps of the proposed reduction procedure: the overall error of the algorithm never exceeds those of each of its subprocedures.

Finally, the absolute and compensated signals previously analyzed have different harmonic contents (typically first harmonic in the absolute and higher order harmonics in the compensated), thus the algorithm's performance could be further improved by means of a separate reduction in the angular domain of such signals.

Acknowledgments

This work is supported by CERN through the agreement LHC/AT/K1464 with the University of Sannio, whose supports authors gratefully acknowledge. Authors thank Felice Cennamo and Louis Walckiers for their useful suggestions.

References

- [1] S. Amet, L. Bottura, L. Deniau, and L. Walckiers, "The multipoles factory: an element of the LHC control", *IEEE Trans. on Appl. Supercond.*, Vol. 12, pp. 1417-1421, 2002.
- [2] N.R. Brooks L. Bottura, J.G. Perez, O. Dunkel, "Estimation of Mechanical Vibrations of LHC Fast Magnetic Measurement System", *Proc. of 20th Biennial Conf. on Mag. Tech. (MT-20)*, Philadelphia, PE, August 27-31, 2007.
- [3] M. Haverkamp, L. Bottura, E. Benedico, S. Sanfilippo, B. ten Haken, H. H. J. ten Kate, "Field Decay and Snapback Measurements Using a Fast Hall Plate Detector", *IEEE Trans. on Applied Superconductivity*, Vol. 12, No. 1, March 2002, pp. 86-89.
- [4] P. Arpaia, A. Masi, G. Spiezia, "A digital integrator for fast and accurate measurement of magnetic flux by rotating coils", *IEEE Trans. on Instrum. Meas.*, Vol. 56, No. 2, April 2007.
- [5] D.A. Huffman, "A Method for the Construction of Minimum-Redundancy Codes", *Proceedings of the I.R.E.*, September 1952, pp 1098-1102.
- [6] Ian H. Witten, Radford M. Neal, and John G. Cleary, *Arithmetic Coding for Data Compression*, CACM 30(6):520-540, June 1987.
- [7] W. B. Pennebaker and J. L. Mitchell, *JPEG Still Image Data Compression Standard*. New York: Van Nostrand Reinhold, 1993.
- [8] Mehmed Kantardzie, *Data Mining: Concepts, Models, Methods, and Algorithms*, IEEE Press, New York, NY, 2003, 360pp., ISBN: 0-471-22852-4.
- [9] K.H. Lee, H. Woo, T. Suk, *Data Reduction Methods for Reverse Engineering*, The International Journal of Advanced Manufacturing Technology, Issue Volume 17, Number 10 / May, 2001, pp. 735-743.
- [10] W. J. Tompkins, *Biomedical digital signal processing*, Prentice Hall, New Jersey, 2000.

- [11] W. C. Mueller, Arrhythmia detection program for an ambulatory ECG monitor. *Biomed.Sci. Instrument.*, 14: 81–85, 1978.
- [12] J. R. Cox, F. M. Nolle, H. A. Fozzard, G. C. Jr. Oliver, AZTEC: a preprocessing program for real-time ECG rhythm analysis. *IEEE Trans. Biomed. Eng.*, BME-15: 128–29, 1968.
- [13] L. N. Bohs, R. C. Barr, Prototype for real-time adaptive sampling using the Fan algorithm, *Med. & Biol. Eng. & Comput.*, 26: 574–83, 1988.
- [14] Qinghua Hu, Daren Yu and Zongxia Xie, Information-preserving hybrid data reduction based on fuzzy-rough techniques, *Pattern Recognition Letters*, Volume 27, Issue 5, 1 April 2006, Pages 414-423.
- [15] J.R. Cano, F. Herrera, M. Lozano, Using evolutionary algorithms as instance selection for data reduction in KDD: an experimental study, *Evolutionary Computation*, IEEE Transactions on, Dec. 2003, Volume: 7, Issue: 6, pag. 561- 575.
- [16] A. Djafari Marbini, L. E. Sacks, “Adaptive sampling mechanisms in sensor networks”, *London Communications Symposium*, London, 2003.
- [17] A. Jain, E. Y. Chang, “Adaptive sampling for sensor networks”, *Proc. the Fisrt Workshop on Data Management for Sensor Networks (DMSN 2004)*, Toronto, Canada, August 30th, 2004.
- [18] J. Kho, A. Rogers, N.R. Jennings, “Decentralized adaptive sampling of wireless sensor networks”, *First International Workshop on Agent Technology for Sensor Networks, a workshop of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-07)*, Honolulu, Hawaii, USA, May 14-18 2007.
- [19] C. Alippi, G. Anastasi, C. Galperti, F. Mancini, M. Roveri, “Adaptive sampling for energy conservation in sensor networks for snow monitoring applications”, *IEEE Internatonal Conference on Mobile Adhoc and Sensor Systems*, 2007. MASS 2007. , 8-11 Oct. 2007 Page(s):1 – 6.
- [20] C. Alippi, M. Roveri, An adaptive CUSUM-based Test for Signal Change Detection, *Proceedings of the IEEE International Symtposium on Circuits and Systems 2006, ISCAS 2006*.
- [21] O. O. Fadiran, P. Molnár, L. M. Kaplan, “Adaptive sampling via histogram equalization using an active walker model”, *Proc. of the 5th IEEE/ACIS International Conference on Computer and information Science*, 2006
- [22] J. Rigau, M. Feixas, M. Sbert, “Entropy-based adaptive sampling”, *Graphics Interface 2003*, Halifax, Canada, june 2003.

- [23] L. Kok-Fung, "Data Reduction Techniques", in *Biomedical Digital Signal Processing*, Prentice Hall, 2000, pp. 193-215.
- [24] B. K. Natarajan, Filtering Random Noise from Deterministic Signals via Data Compression, IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 43, NO. 11, NOVEMBER 1995, pp. 2595-2605.
- [25] P. G. Spazzini, "Effects of calibration data reduction on a probe performance, *Meas. Sci. Technol.*, Vol. 13, No 9, 2002, pp. 1404-1413.
- [26] Q. Zheng, J. C. Klewicki, "A fast data reduction algorithm for molecular tagging velocimetry: the decoupled spatial correlation technique", *Meas. Sci. Technol.*, Vol. 11, No 9, 2000, pp. 1282-1288.
- [27] "IEEE standard for digitizing waveform recorders", *IEEE Standards Board*, 1994.
- [28] L. Bottura, K. N. Henrichsen, "Field Measurements", *CERN Accelerator School Proceedings*, CERN, September 2004, pp. 118-151.
- [29] A. K. Jain, "Harmonic Coils", *CERN Accelerator School Proceedings*, CERN, April 1997, pp. 175-217.
- [30] L. Bottura, "Standard Analysis Procedures for Field Quality Measurement of the LHC Magnets - Part I: Harmonics", *Internal Note*, 1997.
- [31] D. Wenzel, H. Borsi, E. Gockenbach, Noise Suppression and Data Reduction for Partial Discharge Measurements using Orthogonal Transformations, IEEE International Symposium on Electrical Insulation, Pittsburgh, PA USA, June 5-8, 1994.
- [32] P. Arpaia, L. Bottura, V. Inglese, G. Spiezia, "On-field Validation of the New Platform for Magnetic Measurements at CERN", *Measurement*, Volume 42, Issue 1, January 2009, Pages 97-106, ISSN 0263-2241, DOI: 10.1016/j.measurement.2008.04.006.

LIST OF FIGURE CAPTIONS

Fig. 1. Tracking mechanism of the proposed algorithm in the time domain: observation windows and sampling rate adaptation.

Fig. 2. Tracking mechanism of the proposed algorithm, observed band are highlighted; a) the observed band is monitored, b) significant spectral content appears in the observed band, c) an update of the sampling frequency is triggered and observed band changes accordingly.

Fig. 3. Flow chart of the tracking mechanism.

Fig. 4. Combined approach to data reduction: original signal (dots) and signal reduced by means of *adaptive tracking sampler* (circles) as function of angular position and time.

Fig. 5. Rate-distortion plot [24] and heuristic determination of the noise strength ($\varepsilon = 2 \cdot 10^{-5}$ Vs), corresponding to the maximum of the second derivative.

Fig. 6. Sampling rate step responses for different values of α (the maximum change in the sampling frequency allowed in one step, expressed as a fraction of the current sampling frequency).

Fig. 7. Rotating coils measurement principle.

Fig. 8. Combined approach to data reduction in rotating coils measurements: original flux variations (dots) and flux variations reduced by means of *adaptive tracking sampler* (circles) as function of angular position and time.

Fig. 9. The reference dipole calibration bench at CERN.

Fig. 10. Data flow of FFMM data reduction.

LIST OF TABLE CAPTIONS

Tab. I. Comparison of algorithms' static performance on simulated flux increment signal.

Tab. II. Performance of the Fan algorithm, employed by the *noise-canceller compressor*, for the reduction of nearly linear signals.

Tab. III. Comparison of algorithms' computational burden expressed as execution time required to process the same amount of incoming data (a sine wave of 131072 points): computation performed in MATLAB on a Pentium IV-2.8 GHz processor.

Tab. IV. Compression Ratio and RMS Error for different settings of the *adaptive tracking sampler* (ATS) on sinusoidal data of an LHC superconducting dipole.

Tab. V. Compression Ratio and RMS Error of the *noise-canceller compressor* (NCC) run after the *adaptive tracking sampler* (ATS) on linear data of an LHC superconducting dipole.

Tab. VI. Compression Ratio and RMS Error for different settings of the algorithm on data of a resistive reference dipole at constant current (200 A).

Tab. VII. Compression Ratio and RMS Error for different settings of the algorithm on data of a resistive reference dipole at variable current (ramp from 15 to 200 A at 10 A/s).

Fig. 1. Tracking mechanism of the proposed algorithm in the time domain: observation windows and sampling rate adaptation.

Fig. 2. Tracking mechanism of the proposed algorithm, observed band are highlighted; a) the observed band is monitored, b) significant spectral content appears in the observed band, c) an update of the sampling frequency is triggered and observed band changes accordingly.

Fig. 3. Flow chart of the tracking mechanism.

Fig. 4. Combined approach to data reduction: original signal (dots) and signal reduced by means of *adaptive tracking sampler* (circles) as function of angular position and time.

Fig. 5. Rate-distortion plot [24] and heuristic determination of the noise strength ($\epsilon = 2 \cdot 10^{-5}$ Vs), corresponding to the maximum of the second derivative.

Fig. 6. Sampling rate step responses for different values of α (the maximum change in the sampling frequency allowed in one step, expressed as a fraction of the current sampling frequency).

Fig. 7. Rotating coils measurement principle.

Fig. 8. Combined approach to data reduction in rotating coils measurements: original flux variations (dots) and flux variations reduced by means of *adaptive tracking sampler* (circles) as function of angular position and time.

Fig. 9. The reference dipole calibration bench at CERN.

Fig. 10. Data flow of FFMM data reduction.

Tab. I. Comparison of algorithms' static performance on simulated flux increment signal.

Algorithm	Compression Ratio*	Amplitude (mVs)	SINAD (dB)	SNHR (dB)	-THD (dB)
<i>None</i>	1.00	4.674	54.10	54.83	61.51
<i>Turning Point</i>	1.02	4.673	54.31	54.82	63.10
<i>AZTEC</i>	8.00	4.495	24.63	27.64	26.92
	19.49	4.234	20.76	32.57	20.30
	36.51	4.220	16.98	19.29	17.68
<i>Fan</i>	1.02	4.674	54.10	54.83	61.51
	1.15	4.674	54.10	54.83	61.51
	2.39	4.673	53.58	54.50	59.96
	10.66	4.606	40.41	42.94	43.23
	32.00	3.819	19.25	37.62	18.63
<i>Adapt. Tracking Sampler</i>	2.00	4.674	58.76	63.06	59.98
	3.00	4.673	60.47	64.80	61.71
	4.00	4.673	60.81	66.81	61.07
	5.00	4.671	61.86	68.84	62.07

* size of original data divided by size of compressed data

Tab. II. Performance of the Fan algorithm, employed by the *noise-canceller compressor*, for the reduction of nearly linear signals.

Algorithm		Compression Ratio*	Amplitude (mVs)	SINAD (dB)	SNHR (dB)	-THD (dB)
<i>none</i>	σ_{noise} (Vs)					
	$8.47 \cdot 10^{-6}$	1.00	4.674	54.10	54.83	61.51
<i>Fan</i>	ε (Vs)					
	$1.0 \cdot 10^{-6}$	1.28	4.674	54.11	54.84	61.51
	$2.0 \cdot 10^{-6}$	1.55	4.674	54.11	54.84	61.52
	$4.0 \cdot 10^{-6}$	2.15	4.674	54.13	54.87	61.50
	$6.0 \cdot 10^{-6}$	2.90	4.674	54.18	54.93	61.37
	$8.5 \cdot 10^{-6}$	4.25	4.674	54.29	55.06	61.28
	$1.0 \cdot 10^{-5}$	5.48	4.674	54.39	55.16	61.26
	$1.2 \cdot 10^{-5}$	7.71	4.674	54.50	55.32	60.93
	$1.7 \cdot 10^{-5}$	18.5	4.674	54.97	55.85	60.77
	$2.0 \cdot 10^{-5}$	34.5	4.674	54.98	56.04	60.04
	$2.6 \cdot 10^{-5}$	101.2	4.674	55.51	56.83	59.52
	$3.4 \cdot 10^{-5}$	368.1	4.674	55.96	56.86	61.95
	$8.0 \cdot 10^{-5}$	512.0	4.674	55.41	56.51	61.24
	$1.0 \cdot 10^{-4}$	512.0	4.674	55.41	56.51	61.24

* size of original data divided by size of compressed data

Tab. III. Comparison of algorithms' computational burden expressed as execution time required to process the same amount of incoming data (a sine wave of 131072 points): computation performed in MATLAB on a Pentium IV-2.8 GHz processor.

Algorithm	Compression Ratio*	Execution Time (s)
<i>Turning Point</i>	2	72.54
<i>AZTEC</i>	5.39	13.56
<i>Fan</i>	5.34	22.67
<i>Adapt. Tracking Sampler</i>	-.**	24.06

* size of original data divided by size of compressed data

**the time to compute the *fft* does not change with the achieved CR

Tab. IV. Compression Ratio and RMS Error for different settings of the *adaptive tracking sampler* (ATS) on sinusoidal data of an LHC superconducting dipole.

Original signal	131072 points, 1024 S/s					
Reduced Signal	Optimal Sampling Rate (S/s)	Compression Ratio*	Main Component RMS Error Ampl. (*10 ⁻⁶ T)/Phase (rad)		Harmonics RMS Error (ppm)**	
			ATS	Fan alg.	ATS	Fan alg.
	512	2	3.89 / 0.0	0.21 / 0.0	0.189	0.799
	256	4	7.05 / 0.0	4.78 / 0.0	0.315	3.070
	128	8	10.96 / 0.0	254.89 / 0.0	38.261	72.452

* size of original data divided by size of compressed data

**harmonics up to the 10th are considered

Tab. V. Compression Ratio and RMS Error of the *noise-canceller compressor* (NCC) run after the *adaptive tracking sampler* (ATS) on linear data of an LHC superconducting dipole.

ATS	NCC				
Compression Ratio*	Compression Ratio*		Main Component RMS Error		Harmonics RMS Error (ppm)**
	ABS	COMP	Ampl. (*10 ⁻⁶ T)	Phase (rad)	
2	4.31	4.65	74.24	0.0	0.093
4	3.50	4.19	48.15	0.0	0.064
8	4.92	4.31	37.29	0.0	0.062

* size of original data divided by size of compressed data

**harmonics up to the 10th are considered

Tab. VI. Compression Ratio and RMS Error for different settings of the algorithm on data of a resistive reference dipole at constant current (200 A).

Original signal	4096 poits/turn, 1 rps					
Reduced signal		Compression Ratio* ABS COMP		Main Component RMS Error Ampl.($\cdot 10^{-6}$ T) Phase (rad)		Harmonics RMS Error (ppm)**
	ATS	16	16	1846.96	0.0	2.39
	NCC	3.4	3.7	193.05	0.0	4.79
	proposed alg.	54.4	59.2	1723.00	0.0	2.81

* size of original data divided by size of compressed data

**harmonics up to the 10th are considered

Tab. VII. Compression Ratio and RMS Error for different settings of the algorithm on data of a resistive reference dipole at variable current (ramp from 15 to 200 A at 10 A/s).

Original signal	4096 poits/turn, 1 rps					
Reduced signal		Compression Ratio*		Main Component RMS Error		Harmonics RMS Error
		ABS	COMP	Ampl.($\cdot 10^{-6}$ T)	Phase (rad)	(ppm)**
	ATS	16	16	898.26	0.0	1.42
	NCC	4.6	4.7	326.42	0.0	3.32
	proposed alg.	73.6	75.2	936.78	0.0	2.91

* size of original data divided by size of compressed data

**harmonics up to the 10th are considered