Compression Criteria for Voltage Sags Analysis using Wavelets

GUSTAVO PERES B. DE CASTRO SIGMAR M. DECKMANN ANDRÉ A. FERREIRA
Department of Energy Control and Systems
State University of Campinas
Electrical and Computer Engineering School
BRAZIL
peres@dsce.fee.unicamp.br http://www.dsce.fee.unicamp.br

Abstract: - The use of wavelet decomposition to provide data recording and analysis of Power Quality disturbances depends on establishing an acceptable compromise between high data compression and low energy loss capability, in order to preserve the basic waveform features during signal reconstruction. This paper compares three threshold rules to evaluate compression efficiency, using the retained energy indicator as a signal reconstruction quality index.

Key-Words: - Data Compression, Wavelets, Retained Energy Factor, Threshold, Voltage Sags, Power Quality.

1 Introduction

By definition, a voltage *swell* or *sag* occurs if the effective (rms) voltage exceeds the $\pm 10\%$ tolerance range about the nominal voltage for not longer than 1 minute. These voltage amplitude variations are commonly defined by the couple of parameters: *magnitude* and *duration* of the event. The magnitude expresses the extreme value (minimum or maximum) of the residual rms voltage, while the duration indicates the total time the voltage remained out of the tolerance range.

Considering the occurrence frequency of such events, the mere registration of the rough voltage and current waveforms during the disturbance, may result in megabytes of information, representing considerable data storage needs and high transmission costs. To reduce the number of stored data, the cycle rms voltage values may be calculated during the event. But, due to the rms averaging effect, the detailed disturbance fast evolution may be lost. This means the original event's signature may be altered, thus creating some difficulty to identify the disturbing causes and their origin.

Wavelet theory has been used successfully since 1997 as an effective loss compression method. This method, in addition to other lossless compression methods (LZW, Huffman, Arithmetic Encoding), has significantly increased compression efficiency of Power Quality related disturbances, as shown in [1, 2]. Nevertheless, in those references, the threshold criterion for disregarding some decomposition coefficients are not really discussed, they simply use 10% of the maximum absolute value as the disregarding threshold, at each decomposition level.

Moreover, the reconstructed power system waveforms usually are analyzed using signal loss indicators, some of them derived from the audio signal processing methods, such as signal-noise ratio (SNR), the signal compression ratio (CR) and the error measure expressed by the percent rms difference (PRD). Comparisons among power quality indices before and after compression have so far not been used.

In that sense, this paper aims to establish a criterion based on the percentile of retained energy that enables increased compression rates, without compromising the information required to characterize the disturbance. Firstly, the results are evaluated through the usual signal loss indicators (PRD, SNR) and compared with the methodology used in the classic case [1]. Additionally, the reconstructed signals are analysed through power quality indicators for characterization of the voltage sag.

All signals used in this article may be obtained at the IEEE Project Group P1159.2 website, referenced in [3].

2 Motivation and Basic Concepts

The motivation to use wavelets for the compression of power system voltages and currents, which are corrupted by network disturbances, relies on the fact that most of those time located events present non characteristic low-frequency spectrum, and thus concentrate the relevant information in a small number of coefficients. Thus wavelets provide all the needed time-frequency decomposition techniques.

The most appropriate analysis approach for this class of power signals is known as multi-resolution analysis, which decomposes the signal into different frequency ranges, enabling the encoding of relevant coefficients of the original signal.

An excellent reference about the connection between wavelets, multi-resolution analysis and filter banks theories, is presented in [4]. In this reference, the discrete wavelet transform of a signal is calculated via filtering, followed by downsampling, and can be implemented through a two-channel filter bank system. The structure, by which the signal is broken successively into an increasing number of levels, is known as a dyadic tree. The inverse discrete wavelet transform is equivalent to the synthesis operation.

2.1 Influence of Filter Type in Signal Decomposition

Within the universe of voltage sag/swell related disturbances there is infinity of waveforms and, therefore, some filters obviously behave better for one type of event than for others. For practical purposes, however, it is most important to use filters that represent the best compromise between high compression rates and low loss reconstruction capability.

Most applications of discrete wavelet transform to Power Quality problems do not present detailed filter selection criteria for comparison purposes. In general, Daubechies wavelets (dbN)[1] are the most employed, especially db*4* and db*6*. In [5], a MDL (*Minimum Description Length*) algorithm is used to optimize compression efficiency, in which wavelet sym*7* showed the best results for a given signal set.

All filters mentioned above present characteristics of *compact support* (time location) and *vanishing moments* (zeroes associated to fixed shaped signals).

Both characteristics are important to improve the compression process. The first, by the fact that this kind of function allows a good location, both in time and frequency; and the second because it is a property that associates zeros to the signal areas which are constant, linear, quadratic, etc. The number of vanishing moments for these wavelet functions (db*N*) equals *N*.

Due to the appliance on three-phase voltages, and considering that the events may present diverse behaviors between phases, the adopted filter can produce different amounts of coefficients per phase. Although this is not a real problem to handle with, it may be convenient to save a similar number of coefficients per phase solely to provide uniform vector size storage. In this case the smallest values of the longest coefficients vector may be disregarded. Therefore it becomes important to compare the quality index resulting before and after the coefficient reduction.

[1]Where N is the order and $2N$ the size of the filter.

2.2 Selection of Levels Number

Regarding the decomposition levels number, the selection is frequently based on the cut-off frequency of the low-pass filter employed. According to [6, 7], the cut-off frequency for a corresponding j level is:

$$f_{cutoff} = \frac{f_S}{2^{j+1}};$$

where f_S is the sampling frequency.

Hence, for the decomposition into sub-bands, every new stage or level, has its bandwidth reduced by half, due to the decimation process (2^{j+1}). However, at the same time, the signal size is also reduced proportionally.

Considering k as the total number of samples of the input signal, and assuming the filter size is $2N$, the filter convolution with input signal, at each level, produces a new signal with $k + 2N - 1$ samples per branch for a two-channel filter bank. Since the signal is downsampled, the number of coefficients at this stage will be:

$$SIV\left(\frac{k-1}{2}\right) + N. \tag{1}$$

The SIV function returns the *smallest integral value* resulting from the expression$(k-1)/2$.

The compression algorithm adopted in this paper uses all *approximation coefficients*[2] (low-frequency range) of the highest decomposition level (defined as **base signal** according to a lost energy criterion), along with the larger magnitude components of the *detail coefficients*[3] (selected according to a threshold criterion) and their corresponding positions in the data vector. The objective is to reconstruct the signal that reproduces the disturbing event in a reliable way and at the same time allowing the identification of its possible causes. Therefore, a correct selection of the number of decomposition levels is very important to ensure high compression without loss of essential features of the event.

[2]Coefficient set that represents a coarse signal approximation, or low frequency representation.

[3]Coefficient set associated to high signal frequencies with good location in time.

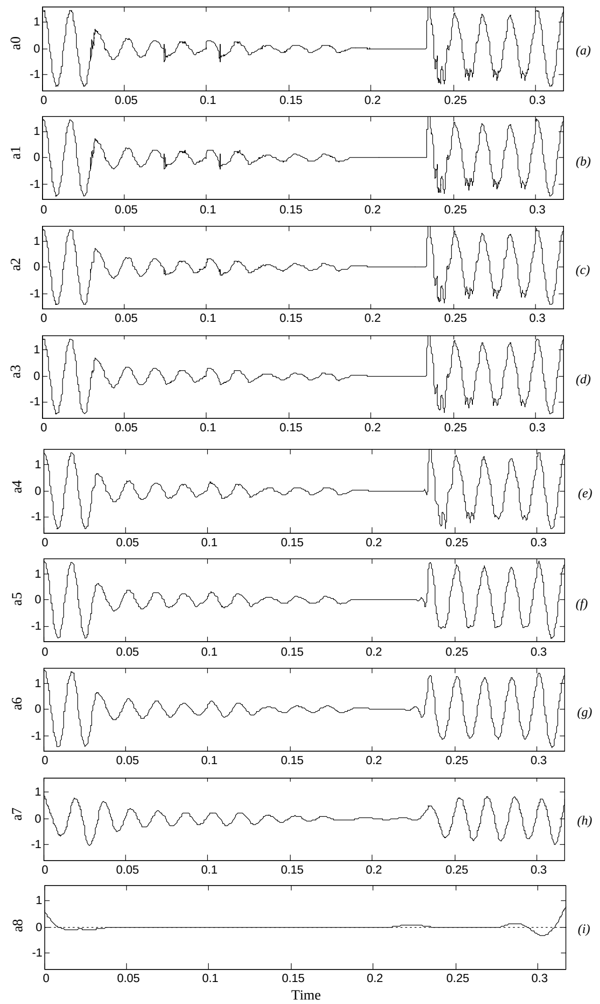

Fig. 1. Approximation signals by decompositions levels using wavelet filter sym*7*: (a) original signal; (b) level 1; (c) level 2; (d) level 3; (e) level 4; (f) level 5; (g) level 6; (h) level 7 and (i) level 8.

Fig.1-(a) presents a voltage sag with 4864 samples. By definition, a_0 are the approximation coefficients (low-frequency) of level zero, which correspond to the actually analyzed signal. From (1), we conclude that the number of approximation coefficients retained at each level in Fig.1 are: (a) 4864, (b) 2438, (c) 1225, (d) 619, (e) 316, (f) 164, (g) 88, (h) 50 and (i) 31; which correspond to compression rates of 49.88%; 74.82%; 87.27%; 93.50%; 96.63%; 98.2%; 98.97% and 99.36% for those coefficients in levels from 1 through 8. Notice that the eighth approximation level (a_8) contains practically no information about the original signal, thus being easily disregarded from the analysis to determine which level presents the best approximation levels to be used as base signal for compression. For this analysis the mother-wavelet employed was sym*7*.

2.3 Threshold Rules

Three rules have been taken into consideration in order to obtain the range of coefficients to be disregarded, according to the hard threshold function of parameter λ calculated as follows:

$$\hat{d}_j = \begin{cases} d_j\left(k^j\right) & \left|d_j\left(k^j\right)\right| \geq \lambda, \\ 0 & \left|d_j\left(k^j\right)\right| < \lambda; \end{cases}$$

where $\hat{d}_j$ are the new detail coefficients and k^j represents the signal size on level (j).

The first threshold rule $(r1)$ is the most traditional, and consists in adopting a percentile of the maximum value per level, e.g. $\lambda^{(r1)} = 0.1 \times \max\left|d_j\left(k^j\right)\right|$.

The second rule $(r2)$ is known as (*VisuShrink*) and has been **adapted** from reference [8] to deal with signals in per unit. It depends on the numbers of samples of the signal, where $\lambda^{(r2)} = \sqrt{0.02\log(k)}$; and k is the total number of samples of the original signal.

The last rule $(r3)$ is a combination of the previous two: $\lambda^{(r3)} = \sqrt{0.02\log(k)} \times \max\left|d_j\left(k^j\right)\right|$, which provides a higher retention amount of the detail coefficients than *VisuShrink* alone, because the term $\max\left|d_j\left(k^j\right)\right| < 1.0$ pu, leads to a lower threshold level ($\lambda^{(r1)} < \lambda^{(r3)} < \lambda^{(r2)}$).

2.4 Compression Efficiency Indicators

For comparison purpose with the method introduced in [1], the first performance evaluation with the above threshold rules was calculated using:

i. The signal-noise ratio (SNR). This index represents the degradation level of the reconstructed signal and is given by:

$$SNR\,(dB) = 10\log_{10}\left\{\frac{\sum\left[a_0\,(k)\right]^2}{\sum\left[a_0\,(k) - \hat{a}_0\,(k)\right]^2}\right\},$$

where $a_0\,(k)$ is the original signal and $\hat{a}_0\,(k)$ is the reconstructed signal with loss.

ii. The compression ratio (CR) defined as:

$$CR\,(\%) = \left\{1 - \left(\frac{T_{compress}}{T_{original}}\right)\right\} \times 100\%,$$

where $T_{compress}$ represents the length of the compressed file and $T_{original}$ the length of the original file.

iii. The percentual rms difference (PRD) as an error measure, given by:

$$PRD = 100 \times \left\{ \sqrt{\frac{\left[a_0\left(k\right) - \hat{a}_0\left(k\right)\right]^2}{\left[a_0\left(k\right)\right]^2}} \right\}.$$

In addition, a decomposition level criterion, based on energy retention, is also considered. The retained energy factor (2) represents a very important compression index for power system disturbances because it expresses the relative energy content of the reconstructed signal.

$$ER\left(\%\right) = 100 \times \left\{ \frac{\left\|\hat{a}_0\left(k\right)\right\|}{\left\|a_0\left(k\right)\right\|} \right\}^2. \qquad (2)$$

It is also possible to define a lost energy criterion (EL) to express the percentual loss of energy of the approximation coefficients (a_j) in each decomposition level (j) .

$$EL_j\left(\%\right) = 100 \times \left\{1 - \left[\left\|a_j\left(k\right)\right\| / \left\|a_0\left(k\right)\right\|\right]^2\right\}.$$

The importance of this lost energy criterion may be pointed out from the flicker content of a voltage amplitude modulation. It is known that a voltage fluctuation of only 0.25% amplitude, at 10Hz modulation frequency produces visible lamp flicker. Thus, by disregarding only 1% of the total energy may eliminate important disturbing components from the point of view of power quality.

However, the main motivation for the use of this criterion is to characterize voltage sags and swells, which are accomplished through the temporal evolution of the voltage rms values. By definition, the rms value represents the constant voltage that produces the same power dissipation on a resistor during the same time interval. Thus *voltage sags/swells are characterized according to the lack or excess of nominal electric energy*.

For decomposition of Fig.1-(a) into approximation coefficients levels 1 through 8, the energy retention according to (2) is shown in Table 1.

Table 1: Energy retention vs. Number of approximation coefficients per level of decomposition.

Level	Number of Coeficients	ER (%)
1	2438	99.99
2	1225	99.89
3	619	99.85
4	316	99.65
5	164	98.58
6	88	97.38
7	50	42.03
8	31	3.35

If, for example, the 1% lost energy criterion is assumed, it is necessary to proceed the decomposition up to level five and store the approximation coefficients of the fourth decomposition level, thus retaining 316 coefficients instead of the 4864 original values. Although the compression rate is 93.5%, the energy retention represents 99.65%.

This decomposition level criterion based on the lost energy has the advantage of not requiring the identification of the most energetic components among the detail coefficients, thus saving computer calculations and processing time.

To guarantee that the reconstructed signal keeps at least 99% of the energy contained in the original signal, a 1.5% lost energy criterion for the coarse signal approximation (**base signal**) was adopted .

3 Test Signals and Results

The signals of Fig.2-(a), with 256 samples per 60Hz cycle, represent an unsymmetrical three-phase short-circuit. Fig.2-(b) shows the corresponding rms voltages evolution, using a half cycle calculation window, with a single sample-moving step.

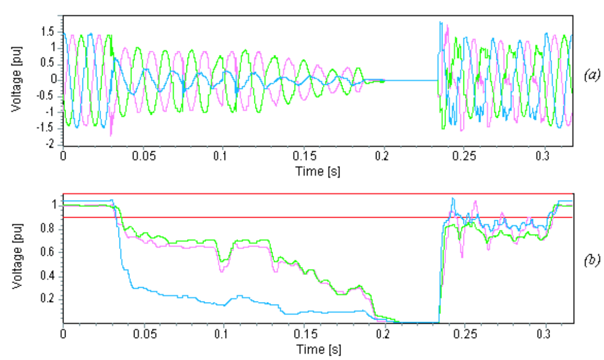

Fig. 2. (a) test voltage; (b) rms voltage.

Following results correspond to decomposition level $j = 5$, using wavelet functions type sym*7* filters. The next tables present the indicators calculated for each phase, according to the threshold rules defined previously.

Table 2: Results for Rule ($r1$).

Fase	CR (%)	SNR (dB)	PRD (%)	ER (%)
A	82.77	22.79	7.26	99.46
B	82.77	35.13	1.75	99.97
C	82.77	22.64	7.38	99.46

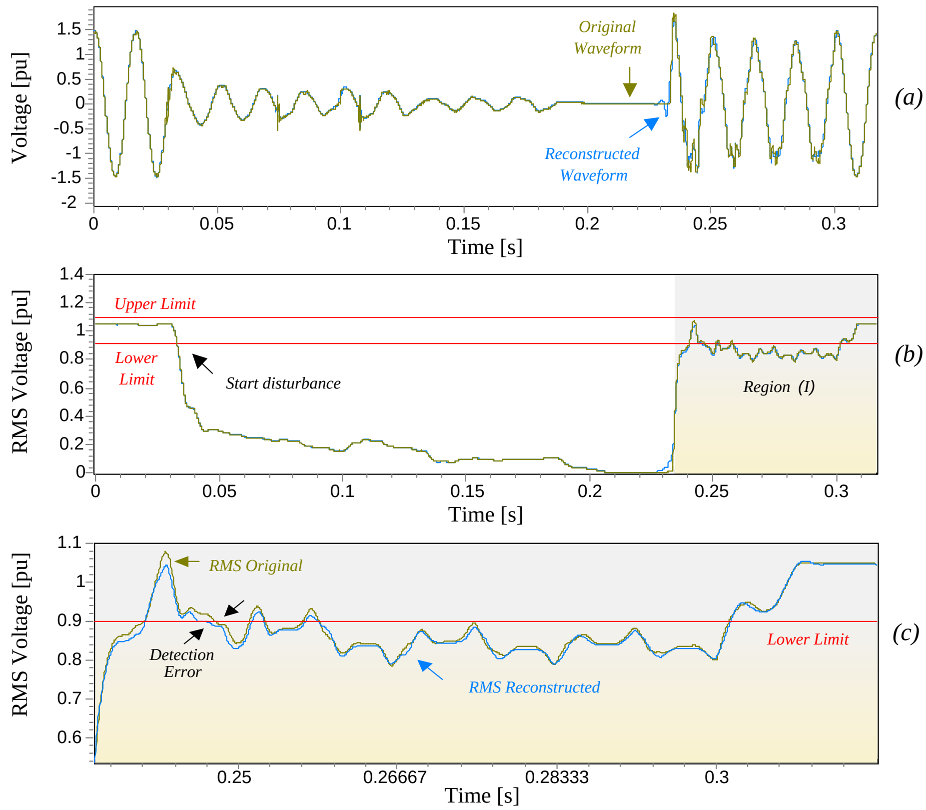

Fig. 3. Original and reconstructed voltage: (a) phase C; (b) rms; (c) amplification of region (I).

Table 3: Results for Rule ($r2$).

Fase	CR (%)	SNR (dB)	PRD (%)	ER (%)
A	95.48	25.45	5.34	99.68
B	95.48	23.56	6.64	99.56
C	95.48	23.28	7.69	99.41

Table 4: Results for Rule ($r3$).

Fase	CR (%)	SNR (dB)	PRD (%)	ER (%)
A	94.16	25.31	5.43	99.67
B	94.16	22.64	7.38	99.45
C	94.16	21.33	8.58	99.26

For comparison, Table 5 presents the results obtained with the decomposition procedure of reference [1], combined with three decomposition levels using filter db*4* and threshold rule ($r1$).

Table 5: Reference Model.

Fase	CR (%)	SNR (dB)	PRD (%)	ER (%)
A	76.56	32.43	2.39	99.94
B	76.56	34.94	1.79	99.97
C	76.56	41.13	0.89	99.99

Note that with Santoso's model the compression rate has been sacrificed in order to retain the coefficients necessary to minimize the rms error of the PRD index. Due to this criterion the energy retention factors (ER) resulted very close to unity.

3.1 Voltage Sag/Swell Index Verification on the Reconstructed Signal

In order to look closer the loss imposed by the threshold rules used for signal compression, a special analysis was considered for phase C. According to Tables 2, 3 and 4, this phase presented the largest PRD (rms difference) and the lowest ER (energy retention) among the phases, thus representing the worst case for rules ($r1$), ($r2$) and ($r3$).

Fig.3-(a) presents the original and the reconstructed waveforms of phase C. Also the rms values are shown in Fig.3-(b). It is worth noting that the essential waveform features have been preserved, even during the disturbing event. This waveform similarity is important, for instance, to verify the correct action/settings of the protection system.

For the above analysis of phase C, only threshold rule ($r3$) was chosen. This option was based on the good performance observed using short-term waveforms (6 cycles, 1536 samples records), obtained from [3]. For those cases the mean compression rate (CR) resulted close to 90%, the rms differences

(PRD) below 10% and the retained energy (ER) factors above 99%.

Region (I) of Fig.3-(b) is the most critical for analysis, since it presents several violation points of the voltage tolerance range. Therefore a special look to this region is necessary. Note in Fig.3-(c) that the rms voltage of the reconstructed signal follows very closely the original rms values. This means that the event impact and classification may be correctly evaluated from the reconstructed signal, using the proposed compression technique.

Concerning the sag magnitude and duration detection errors using the reconstructed waveform they resulted, respectively $\approx 0\%$ and ≈ 0.1 cycle, according to details shown in Fig.4.

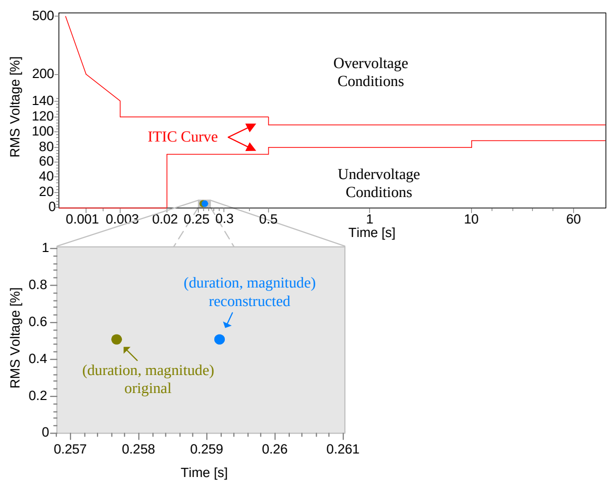

Fig. 4. ITIC acceptance curve: magnitude and duration against the ITIC limits.

4 Conclusion

Voltage sags and swells constitute a special class of unpredictable disturbing events. A partial voltage sag may be caused by a short-circuit occurring far away from the observation point. The propagation of such events along the power system is not well defined and thus very difficult to predict. This also rises considerable difficulties to get consensus about the best way to characterize those disturbances.

It is thus important to save the more information possible concerning voltage and current disturbances for post analysis to be performed in the interest of utilities, industry, customers and standardization agencies.

The present work proposes a data compression procedure via discrete wavelet transform with low computational burden, with a flexible threshold rule, adaptable for different signal sizes, and resulting more conservative for short time record signals.

The retained energy (ER) factor was adopted in order to guarantee high compression rates and to obtain high quality reconstruction performance for voltage sags and swells off-line analysis.

The essence of the compression criterion for power systems is the maximum energy retention for the low frequency approximation coefficients. This includes the fundamental and harmonic content of the power signals. The high frequency components are typically associated with time located fast transients. They are therefore characterized by the detail coefficients, presenting lower energy content.

Acknowledgement:
The authors wish to express their gratitude to CNPq (the National Council for Scientific and Technological Development) for providing the financial support for this research.

References:

[1] S. Santoso, E. J. Powers, and W. M. Grady, "Power quality disturbances data compression using wavelet transform methods," *IEEE Transactions on Power Delivery*, vol. 12, no. 3, pp. 1250–1257, March 1997.

[2] T. B. Littler and D. J. Morrow, "Wavelets for the analysis and compression of power system disturbances," *IEEE Transactions on Power Delivery*, vol. 14, no. 2, pp. 358–364, April 1999.

[3] IEEE P1159.2, "Monitoring electric power quality - ieee std. 1159," 1995, IEEE Task Force - Project Group P1159.2. Disponível em: http://grouper.ieee.org/groups/index.html.

[4] G. Strang and T. Q. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley, EUA, 1996.

[5] E. Y. Hamid, Z. I. Kawasaki, H. Yoshida, and H. Doi, "Wavelet and wavelet packet data compression of power system disturbances," *Transactions of the Institute of Electrical Engineers of Japan*, vol. 121, no. 9, pp. 1221–1227, July 2001.

[6] I. Y. Gu and M. H. J. Bollen, "Time-frequency and time-scale domain analysis of voltage disturbances," *IEEE Transactions on Power Delivery*, vol. 15, no. 4, pp. 1278–1284, October 2000.

[7] D. Borrás, M. Castilha, N. Moreno, and J. C. Montaño, "Wavelet and neural structure: A new tool for diagnostic of power system disturbances," *IEEE Transactions on Industry Applications*, vol. 37, no. 1, pp. 184–190, January 2001.

[8] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, April 1994.