High resolution parametric description of slow wave sleep

Piotr J. Durka [a,1], Urszula Malinowska [a,1], Waldemar Szelenberger [b], Andrzej Wakarow [b], Katarzyna J. Blinowska [a,1]

[a] *Laboratory of Medical Physics, Institute of Experimental Physics, Warsaw University, ul. Hoża 69, 00-681 Warszawa, Poland, http://brain.fuw.edu.pl.*

[b] *Medical University of Warsaw, Department of Psychiatry, ul. Nowowiejska 27, 00-665 Warszawa, Poland*

Abstract

We propose a new framework for quantitative analysis of sleep EEG, compatible with the traditional analysis, based upon adaptive time-frequency approximation of signals. Using high resolution description of EEG rhythms and transients in terms of their time occurrence and width, frequency and amplitude, we present a detailed detection and parameterization of slow waves, including also the time occupied by each slow wave—a parameter inaccesible directly by previously applied signal processing methods.

To validate the proposed parameterization, we construct a simple detector of sleep stages 3 and 4, based explicitly upon the classical criteria related to slow waves. To properly compare its performance to the inter-expert agreements and other expert systems, we sort out and discuss the methodology of reporting concordance in this context.

Since the proposed parameterization proves to be compatible with the visual analysis of EEG, we can derive new variables for quantitative analysis of EEG patterns recognized for decades. As examples, we present a continuous description of slow waves and sleep spindles in the overnight sleep, and compare results to the traditional FFT-based estimates.

Key words: EEG slow waves, delta waves, SWA, sleep spindles, matching pursuit, time-frequency, hypnogram, contingency tables, reporting concordance

1 Introduction

The most widely used interpretation of polysomnographic recordings relies on division into sleep stages conforming to the standard criteria, summarized by Rechtschaffen and Kales (1968) (R&K). These rules were originally intended as a reference method; nevertheless, for over 35 years they have being used as a gold standard. This naturally causes severe problems, especially for describing sleep of elderly (Larsen et al., 1995), patients with epilepsy (Marzec and Malow, 2003), and sleep disordered breathing (Norman et al., 2000). Also, stages described by R&K do not reflect many relevant features of sleep EEG. Among the most needed extensions, Billiard (2000) quotes:

- finer time scale than the arbitrary division into 20–30 sec. epochs,
- a measure of spindle intensity,
- quantitative determination of EEG components in the low frequency range.

We propose a unified solution to these issues by means of a repeatable procedure, based upon explicit parameterization of local EEG patterns. Waveforms present in the signal are described not only in terms of their frequency and amplitude (energy), but, contrary to the previously applied methods, their exact time positions and durations are also determined. This approach conserves compatibility with the visual EEG analysis.

To test the automatic parameterization of slow waves, in this paper we evaluate detection of the deep sleep stages (3 and 4) based directly upon the classical R&K (Rechtschaffen and Kales, 1968) criteria. Results are compared to visual detection and scoring, including the inter-expert concordance. Different ways of reporting concordance of detection are summarized and discussed.

Robust and high-resolution automatic detection and parameterization of sleep EEG patterns, compatible with visual EEG analysis, opens new possibilities of computer aided analysis. In this paper we present examples of new descriptors of sleep spindles and slow waves, which are more sensitive than the spectral estimates and fully compatible with the traditional approach.

Email addresses: durka@fuw.edu.pl (Piotr J. Durka), ula@fuw.edu.pl (Urszula Malinowska), wald@psych.waw.pl (Waldemar Szelenberger), awakar@psych.waw.pl (Andrzej Wakarow), kjbli@fuw.edu.pl (Katarzyna J. Blinowska).

[1] Partially supported by Committee for Scientific Research (Poland).

2 Materials and Methods

2.1 Experimental data

19 healthy subjects were included in the study. Polysomnograms were recorded during two consecutive nights, only the data from the second night were analyzed. Data were acquired from 23 channels including standard 10/20 derivations, FPz, Oz, A1 and A2 in reference to an electrode placed between Cz and Fz. EOG and EMG were recorded according to Rechtschaffen and Kales (1968) criteria.

The signal was filtered with an analog bandpass filter (0.15-30 Hz) and then sampled with frequency 128 Hz. Analog-digital 12-bit converter was used. Silver electrodes were applied with collodion. Maximal resistance was 5 kΩ.

Visual analysis of the sleep EEG, based on 20 sec. epochs, was performed from standard polysomnographic derivations, according to Rechtschaffen and Kales (1968) rules, by experienced sleep researcher. Artifacts (ocular, muscle, technical, related to breathing, ECG and electrode contacts) were visually marked in fixed 4-sec epochs in each polysomnogram. Automatic analysis presented in this paper refers to the EEG derivation C3–A2 from sleep onset to final awakening. Concordance of scoring was calculated on artifact-free epochs, Figures 2–3 contain also epochs marked as containing artifacts.

Informed consent was obtained from all the subjects. The study was approved by the University Ethics Committee.

2.2 Matching pursuit

Matching pursuit is an algorithm introduced in Mallat and Zhang (1993) as a sub-optimal solution to the intractable problem of an optimal approximation of a signal (f) in a redundant dictionary of functions (D). In the first step of MP, the waveform g_{γ_0} which best matches the signal f is chosen from the dictionary D. In each of the consecutive steps, the waveform g_{γ_n} is matched to the signal $R^n f$, which is the residual left after subtracting results of previous iterations:

$$\begin{cases} R^0 f = f \\ R^n f = \langle R^n f, g_{\gamma_n} \rangle g_{\gamma_n} + R^{n+1} f \\ g_{\gamma_n} = \arg\max_{g_{\gamma_i} \in D} |\langle R^n f, g_{\gamma_i} \rangle| \end{cases} \tag{1}$$

For a complete dictionary the procedure converges to f in theoretically infinite number of iterations, but in practice relatively few waveforms provide very good approximation:

$$f = \sum_{n=0}^{M} \langle R^n f,\ g_{\gamma_n} \rangle g_{\gamma_n} \tag{2}$$

Functions g_γ are chosen from dictionaries composed—apart from the complete Dirac and Fourier bases—from the Gabor functions, since these functions provide optimal joint time-frequency localization Mallat (1999). Real valued Gabor can be expressed as

$$g_\gamma(t) = K(\gamma) e^{-\pi\left(\frac{t-u}{s}\right)^2} \cos\left(\omega(t-u)+\phi\right), \tag{3}$$

where $K(\gamma)$ is such that $||g_\gamma|| = 1$. Parameters $\gamma = \{u, \omega, s\}$ of the possible Gabor functions (phase ϕ is optimized separately in numerical implementations) constitute a 3-dimensional continuous space, from which a finite dictionary must be chosen for an implementation of the procedure given by Eq. (1). In the algorithm described by Mallat and Zhang (1993), dictionary was composed from atoms chosen from predefined dyadic sequences. However, any fixed scheme of subsampling the space of possible dictionary's functions leads to a statistical bias of the resulting decompositions. A solution proposed in Durka et al. (2001) relies on stochastic dictionaries, in which parameters $\{u, \omega, s\}$ are drawn from uniform distributions across ranges defined by sizes of the signal and the dictionary. Complete software implementation of a bias-free MP decomposition, used in this study, is freely available from `http://eeg.pl/mp`.

2.3 Reporting concordance

Before any reasonable evaluation/comparison of an automatic method with the traditional, human detection, we must face the fact that the very term "concordance" does not seem to have a precise meaning in this context. Different works report inter-expert concordance in sleep scoring (Danker-Hopfe et al. (2004); Kim et al. (1992); Norman et al. (2000)), or concordance between human experts and automatic systems (Ferri et al. (1989); Hae-Jeong et al. (2000); Hashizume et al. (2001); Schaltenbrand et al. (1996)). Concordance is estimated by means of different measures—usually related to sensitivity or selectivity, but in many cases it's impossible to guess what the given percentage of "concordance" may actually mean. Therefore, below we briefly review and discuss ways of reporting agreement in sleep scoring.

The first step in any comparison of scorings of two experts in few categories is creating a contingency table, that is a table of counts that cross-classifies the data. As an example we will use Table 1, related to the inter-expert agreement in scoring stages 3, 4, and the third group, including stage 1, 2, REM, wakefulness and muscle, in this study pooled together and termed "nonSWS". It counts numbers of epochs, classified by the two experts in any of the available categories (sleep stages). Main diagonal of the contingency table contains counts of epochs, in which both experts agreed on the stage score. Off-diagonal cells count disagreement, that is cases when experts assigned different stage score to the same epoch. For example, "10" in the 3rd row of 2nd column means that 10 epochs were scored as stage 4 by Expert A and as stage 3 by Expert B.

	Expert B					
	stage	**nonSWS**	**stage 3**	**stage 4**	**SUM**	**% agr.**
	nonSWS	**2753**	299	55	3112	88%
Expert A	**stage 3**	47	**112**	161	320	35%
	stage 4	1	10	**447**	458	98%
	SUM	2806	421	663	3890	
	% agr.	98%	27%	67%		

Table 1
Example of a contingency table: inter-expert agreement in scoring stages 3 and 4, evaluated on 3 overnight recordings scored by 2 experts. Total concordance 85%, Kappa coefficient 0.63, nonSWS concordance 87%, Stage 3 concordance 18%, Stage 4 concordance 66%, *delta* concordance 64% (see section 3.2).

Column labeled SUM indicates the total number of epochs in each category for expert A, the SUM row—for expert B. The last row indicates percentage of the agreement between experts in each category, with expert B treated as a reference. For example, the percentage agreement for the "nonSWS" category, when expert B is treated as the reference is $\frac{2753}{2806} \times 100\% = 98\%$. It equals the sensitivity of expert A in scoring nonSWS epochs. The last column indicates percentage agreement between experts for each category, with expert A as reference. For example, the percentage agreement for nonSWS, when expert A is treated as the reference is $\frac{2753}{3112} \times 100\% = 88\%$. This is the selectivity of expert A in this category.

Percentage agreement between two scorers in all epochs is calculated by summing the values from the main diagonal of the contingency table, and dividing by the total number of epochs. In addition, the Cohen's Kappa is calculated to access the non-random component of this concordance (Cohen, 1960). From a statistical point of view, Kappa coefficient should be preferred as a control parameter (Danker-Hopfe et al., 2004). It measures the agreement between two scorers that is in excess of what might be expected to occur by chance.

Independently of Kappa, it is also possible to measure the proportion of agreement between two judges within each of the categories separately. In Table 1, experts A and B agreed on a total of 2753 nonSWS epochs. For expert A there were 354 additional items in this category, which B did not agree with, while for expert B there were 48 additional nonSWS epochs which A did not agree with. Therefore, the proportion of agreement for nonSWS is $100\% \times 2753/(2753 + 354 + 48) = 90\%$. Consideration of agreements in each category gives a reliable measure of concordance. Quoting only sensitivity or selectivity of one of the scorers, which is a common practice in reporting concordance, does not provide enough information.

We calculated also the concordance for stages 3 and 4 treated as one category—the delta concordance—according to the above proportions.

3 Results

Using the algorithm described in Section 2.2, 20-sec epochs of sleep EEG from derivation C3-A2 were decomposed into a sum (eq. 2) of basic waveforms g_γ (eq. 3). These functions are characterized, apart from the energy (or amplitude) related to $\langle R^n f,\ g_{\gamma_n}\rangle$, by the set of parameters $\gamma = \{u, \omega, s\}$ (time position, frequency and duration).To select sleep spindles and slow waves, we choose from these functions those conforming to the criteria given in Table 2. Minimum amplitude of a slow wave is discussed in Section 3.1. Frequency range of sleep spindles was set slightly wider than the classical 12–14 Hz, as suggested by our experience, and Amzica and Steriade (1998); Hae-Jeong et al. (2000); Schimicek et al. (1994).

	frequency	time duration	min. amplitude
slow waves	0.2–4 Hz	0.5–∞ s.	$A_{\text{min(SWA)}}$ (Sec. 3.1)
sleep spindles	11–15 Hz	0.5–2.5 s.	15 μV

Table 2
Criteria defining sleep spindles and slow waves, applied in this study.

By applying the above criteria to the MP decomposition of sleep EEG, we achieve an automatic detection and exact parameterization of sleep spindles and slow waves. Contrary to the previous approaches, this parameterization includes also explicitly the length of each of these waveforms.

3.1 Minimum amplitude of slow waves

It is generally acknowledged, that the fixed threshold for the minimum amplitude of slow waves causes problems in any automatic detection because of the differences in EEG amplitudes due to age and sex differences (Danker-Hopfe et al., 2004; Larsen et al., 1995) and quality of sleep (Armitage et al., 2000). Since most of these factors seem to influence the overall amplitude of EEG, we adjusted the cutoff for slow waves to the mean overall EEG amplitude, quantified by the standard deviation of the EEG signal in a given channel. As the "optimal cutoff" we took the values, found separately for each recording via the following procedure:

(1) We consider possible cutoff amplitudes A^i chosen to cover a reasonable range between 20 and 120 μV in small ($2\mu V$) steps.
(2) For each of the A^i, we perform a rudimentary scoring by assigning each epoch, occupied by slow waves of amplitude exceeding A^i from 20% to 50% of time, to stage 3, and above 50%—to stage 4.
(3) As the optimal cutoff we take this A^i, for which a posteriori concordance of the above procedure with visually scored hypnograms is maximum.

Each cross in Figure 1 corresponds to such an optimum value, calculated for one of the analyzed recordings. To these optimal values we fitted a linear relation:

$$A_{\mathrm{min(SW)}} = 0.99 \times A_{\mathrm{EEG}} + 28.18\mu V \tag{4}$$

where A_{EEG} reflects the (mean square root of) EEG energy, calculated as the the standard deviation of all the artifact-free epochs of the EEG channel. As presened in Equation (4), in analyzed recordings the cutoff amplitude of slow waves can be obtained approximately as A_{EEG} plus 30 μV. Equation (4) was used to determine the minimal slow wave amplitude for each recording/channel in further analysis. Obtained values were in most cases lower than $75\mu V$ (Figure 1).

3.2 Concordance of parameterization of slow waves with visual detection and sleep scoring

Detection of stages 3 and 4 seems to be the best way to check the validity of the proposed slow waves parameterization, because:

- Sleep scoring is a standard procedure done routinely, hence a large amount of data is available for comparison.

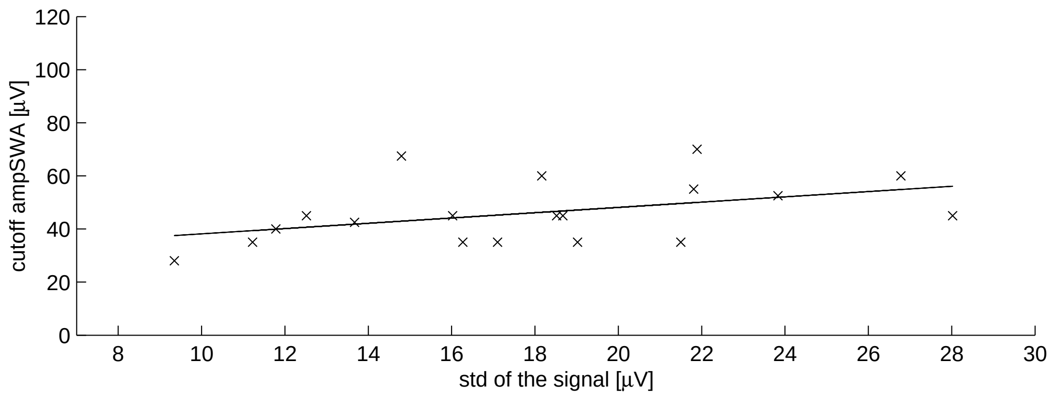

Fig. 1. Optimal cutoff amplitude for slow waves, determined by a posteriori comparison with standard polysomnographic parameters, vs. standard deviation (std) of the signal in entire C3–A2 channel. Each of the 19 recordings of one patient's overnight sleep is represented by a cross. Cutoff amplitude, giving a posteriori maximum concordance of the resulting detection of stages 3 and 4, can be read from the ordinate. On the abscissa standard deviation of the relevant EEG channel, calculated for artifact-free epochs. Straight line (Eq. 4) was fitted by least squares.

- Concordance of a direct marking (tagging) of separate slow waves by an expert and the algorithm is difficult to measure, due to the unclear interpretation of possibly different markings of the start and end of each waveform.

However, when evaluating presented results we must be aware, that in this study for scoring stages 3 and 4 we used only the percentage of delta activity per epoch in a single channel (C3–A2). Actual sleep scoring employs much more of the information, available in a polysomnographic recording. This effect is empirically demonstrated in Table 3. Sleep scorer was asked to mark all the occurrences (start and end) of delta waves in 5 recordings. Using these tags, without any extra information, stages 3 and 4 were mechanically identified via direct application of the R&K criteria: if slow waves occupied more than 50% of epoch's time, it was labelled stage 4, from 20% to 50%—stage 3, below 20%—nonSWS (all other stages). These stages were then compared to the standard hypnogram, evaluated for these recordings by the same expert using a standard procedure—that is employing all the available information, not only the raw percentage of epoch's time occupied by high amplitude slow waves.

Considering all these limitations, we can interpret the concordance of automatic detection of stages 3 and 4 with hypnograms scored by human experts, given in Table 4, as quite satisfactory—that is comparable to the inter-expert concordance reported in literature (Danker-Hopfe et al., 2004; Kim et al., 1992; Monroe, 1967; Norman et al., 2000; Schaltenbrand et al., 1996) and Table 1.

stage	**nonSWS**	**stage 3**	**stage 4**	**sum**	**% agr.**
nonSWS	**4484**	70	35	4593	97%
stage 3	504	**283**	116	903	31%
stage 4	39	125	**676**	840	80%
sum	5031	478	827	6336	
% agr.	89%	59%	82%		

Table 3
Concordance of an mechanical application of just the "20–50%" R&K rules to the slow waves marked by an expert (horizontal) with Stages 3 and 4 from a standard hypnogram by the same expert (vertical). Total concordance 86%, Kappa coefficient 0.65, nonSWS concordance 86%, Stage 3 concordance 26%, Stage 4 concordance 68%, *delta* concordance 65%.

stage	**nonSWS**	**stage 3**	**stage 4**	**SUM**	**% agr.**
nonSWS	**7120**	556	79	7755	92%
stage 3	561	**722**	499	1782	41%
stage 4	130	355	**1205**	1690	71%
SUM	7811	1633	1783	11227	
% agr.	91%	44%	68%		

Table 4
Concordance of the automatic detection of stages 3 and 4, based upon MP parameterization of slow waves taken as waveforms conforming to the time width and frequency from Table 2 and amplitude from Eq. 4 (horizontal), with hypnograms by human expert (vertical). Total concordance 81%, Kappa coefficient 0.59, non-SWS concordance 84%, Stage 3 concordance 27%, Stage 4 concordance 53%, *delta* concordance 68%.

3.3 Parametric description of slow waves and spindles

In the previous section we presented concordance of the automatic parameterization of slow waves, achieved via the MP algorithm, with the visual scoring of slow wave sleep. Concordance of MP-based approach with visual detection of sleep spindles was presented by Żygierewicz et al. (1999). Results presented by Durka et al. (2002) in the field of pharmaco-EEG confirm also general concordance of the MP-based approach with the classical methodology.

Using such a detailed, reliable and sensitive parametrization of slow waves and sleep spindles, we can construct a variety of EEG descriptors, involving relevant properties of these waveforms: energy, amplitudes, occurrences in time etc., as well as any parameters derived from these. As an example, Figure 2 presents the time course of the number of sleep spindles per 20-sec epoch, and the percentage of epoch's time occupied by the waveforms classified as delta waves of high amplitude (both detected according to the procedure described in section 3). Lines denoting the 20% and 50% of epoch's time, occupied by delta waves of high amplitude, correspond to the classical R&K criteria. Classification of epochs exceeding 50% as stage 4, those falling between 20% and 50% as stage 3, and the rest as nonSWS, corresponds to the raw procedure presented in the previous section. Apart from this direct correspondence, we observe also continuous and gradual changes of SWA.

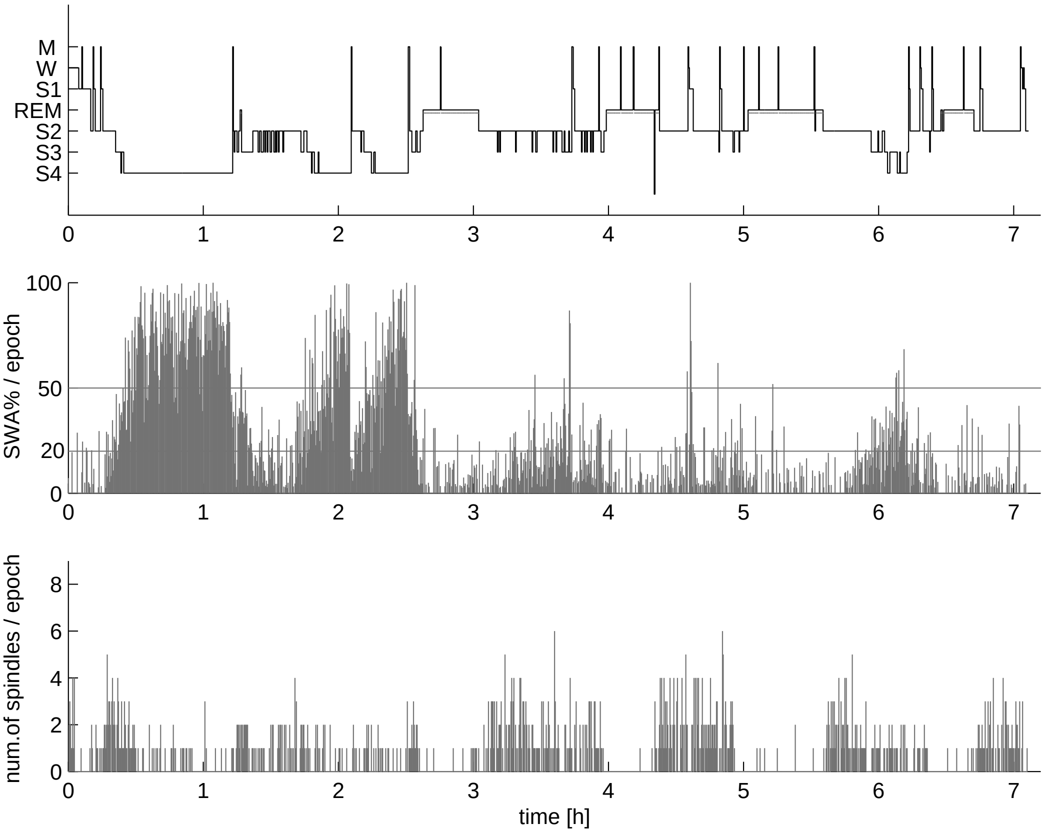

Fig. 2. Top: hypnogram (by human expert); bottom—slow waves and sleep spindles detected automatically: % SWA denotes the percentage of epoch occupied by waveforms classified as slow waves of high amplitude. This continuous description of the slow wave sleep is fully compatible with stages 3/4 classification by Rechtschaffen and Kales (1968), as indicated by the 20% an 50% lines. Artifacts not removed from analysis.

Similar description can be obtained also from the classical, spectral estimates. Figure 3 compares spectral integrals, evaluated in frequency ranges given in Table 2, to the corresponding energy estimates, obtained by summing the energy of relevant waveforms parameterized by the MP algorithm. However, due to the low selectivity of the spectral estimators, they contain a significant background—in both frequency ranges of SWA and sleep spindles—across all the duration of the sleep, not only in stages when these activities actually occur. If we were to interpret these results directly, we would have to say e.g. that there is a certain amount of spindling activity present also across all the REM stages. On the contrary, due to the high resolution and sensitivity of the presented approach, erroneous detections in MP-based estimates are mostly confined to the epochs containing artifacts (in Figures 2 and 3, artifact contaminated epochs were not eliminated to preserve the continuity of time—artifacts amounted to a significant percentage of most of the polysomnographic recordings used in this study). Figure 4 gives a heuristic explanation of the theoretical foundations of this effect.

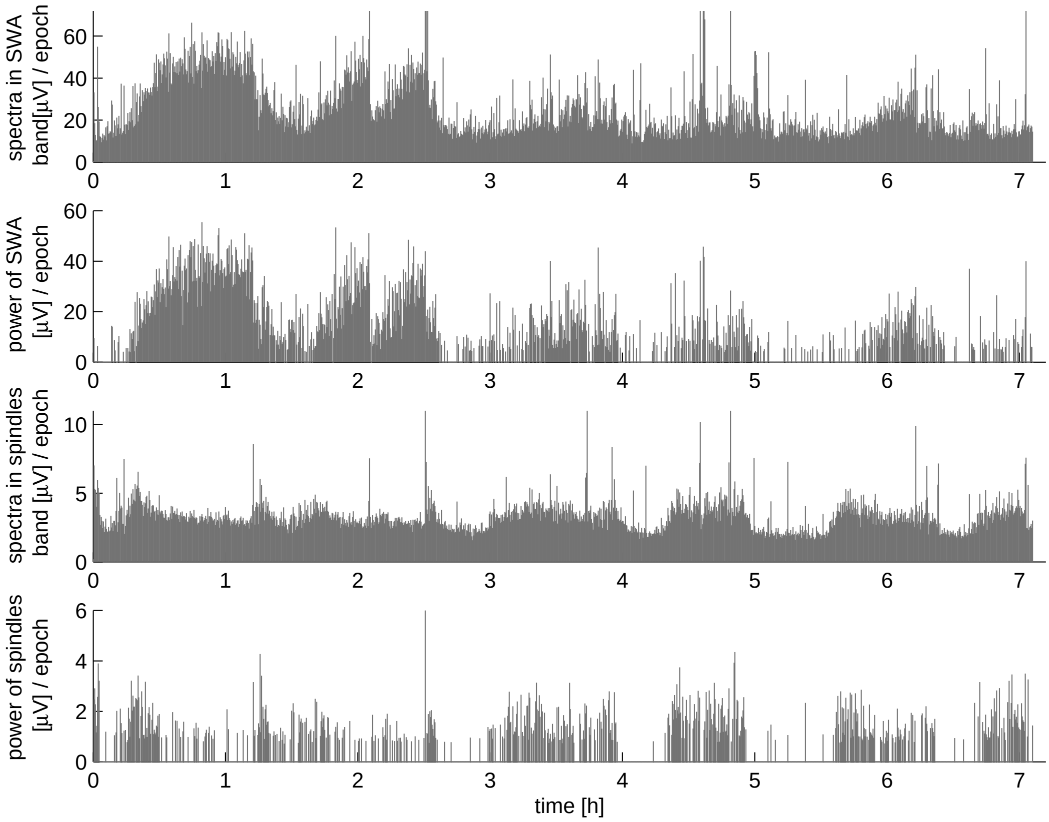

Fig. 3. Top: square root of the power of SWA, calculated from spectral integrals in ranges corresponding to Table 2 and power (sqrt) of SWA estimated for waveforms detected via the proposed procedure. Bottom—the same for sleep spindles. Power in square root scale, time of sleep in hours. As in Fig. 2, artifact contaminated epochs are not removed to preserve the continuity of time. We observe higher sensitivity of the proposed approach, compared to the spectral estimates incorporating significant noise.

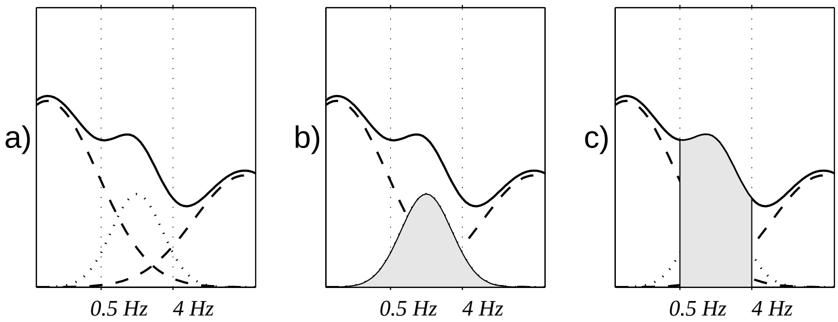

Fig. 4. Explanation of the effect presented in Figure 3—a consequence of the uncertainty principle in signal analysis. Spectral power (vertical) versus frequency (horizontal) of three hypothetical waveforms with frequency centers lying inside (dotted) and outside (dashed line) the frequency interval, assumed in this study for slow waves. Due to the uncertainty principle, their spectral contents overlap. Solid line presents their sum, i.e. total spectral power, as estimated e.g. by Fourier transform. In b) we mark (shaded) the actual power carried by a hypothetical slow wave, as estimated in this study. Plot c) highlights the power obtained from a spectral integral from 0.5 to 4 Hz. We observe that neighboring waveforms from outside the interval of interest may contribute significantly to the activity estimated within the interval, while some of the power carried by the slow wave falls outside the interval and is neglected in the spectral estimate. In the absence of the central structure (slow wave), spectral estimates still reveal some power in the SWA range, which gives the background activity present across all the timeline in the topmost plot of Figure 3.

4 Discussion

This study, and the accompanying software available freely from `http://eeg.pl/mp`, provide foundations for a computer aided analysis of sleep EEG, based upon adaptive time-frequency approximation. Apart from describing the EEG rhythms and transients in terms of their amplitudes and frequencies, it offers also a new parameter, describing their duration explicitly and with high resolution.

To assure compatibility with the traditional analysis, we evaluated this parameter via rudimentary algorithm of scoring sleep stages 3 and 4. In the previous works, inter-expert concordance between 52% and 94% have been obtained for visual scoring of stage 3, and between 44% and 94% in scoring of stage 4, depending on subject groups and whether the comparison were made within or between laboratories (Kim et al., 1992; Martin et al., 1972; Schaltenbrand et al., 1996). Due to the non-standard ways of reporting concordance in most of these works, we cannot reliably compare them to the results obtained in this study; however, exploring detailed information contained in Table 1 suggests that the inter-expert agreement evaluated in this study was on a generally similar level.

Schaltenbrand et al. (1996) reported an automatic staging system achieving sensitivity between 46.7 and 78.1% for stage 3, and from 56.7 to 94.5% for stage 4. Ferri et al. (1989) achieved agreement in detection of delta stages varying from 3% to 100%. Again, the most relevant measures of concordance are not quoted. Nevertheless, we may read similar sensitivities from Table 4: 44% for stage 3 and 68% for stage 4. They suggest that the performance of the system, constructed in this study ad hoc to validate the detection of slow waves, is comparable to other expert systems. Given the limitations of a mechanical detection of deep sleep stages, based solely upon the epoch's percentage occupied by slow waves, obtained results (given in Table 4) confirm the correctness of slow waves parametrization within the proposed framework.

Performance of detection naturally depends on the applied definitions of EEG patterns (Section 3). Strict and direct application of the classical criteria, like e.g. 12-14 Hz frequency range for sleep spindles, does not imply the best concordance with visual detection. Another generally recognized and open problem is the inter-subject variability of the EEG amplitude, affecting directly R&K definition of slow waves (minimum amplitude 75 μV). We found a simple heuristic relation, which for the datasets used in this study related the minimum amplitude of slow waves to the average amplitude of EEG (quantified by standard deviation) plus a constant.

Using the proposed parameterization, we may construct new, continuous and high resolution descriptors of the sleep process, without violating the Occam's parsimony principle. Unlike many new approaches, descriptors based upon the proposed paradigm can be fully compatible with R&K criteria (as presented e.g. in Figure 2), and explainable in well established terms like sleep spindles or slow waves. This should significantly facilitate the acceptation of advanced signal processing methods in clinical environments.

5 Acknowledgments

This work was partly supported by a grant of the State Committee for Scientific Research (KBN) to the Institute of Experimental Physics of Warsaw University.

References

Amzica, F., Steriade, M., 1998. Electrophysiological correlation of sleep delta waves. Electroencephalography and clinical Neurophysiology 107, 69–83.

Armitage, R., Hoffmann, R., Fitch, T., Trivedi, M., Rush, A., 2000. Temporal characteristics of delta activity during nrem sleep in depressed outpatients and healthy adults: group and sex effects. Sleep 23 (5), 607–617.

Billiard, M., 2000. The standardized sleep manual, love it or leave it (guest editorial). Sleep Medicine Reviews 4 (2), 129–130.

Cohen, J., 1960. A coefficient of agreement for nominal scales. Educational and Psychol Meas 20, 37–46.

Danker-Hopfe, H., Kunz, D., Gruber, G., Klosch, G., Lorenzo, J., Himanen, S., Kemp, B., Penzel, T., Roschke, J., Dorn, H., Schlogl, A., Trenker, E., Dorffner, G., 2004. Interrater reliability between scorers from eight european sleep laboratories in subjects with different sleep disorders. Journal of Sleep Research 13 (1), 63–9.

Dijk, D., Beersma, D., Bloem, G., 1989. Sex differences in the sleep eeg of young adults: visual scoring and spectral analysis. Sleep 12 (6), 500–507.

Durka, P., Szelenberger, W., Blinowska, K., Androsiuk, W., Myszka, M., 2002. Adaptive time-frequency parametrization in pharmaco EEG. Journal of Neuroscience Methods 117, 65–71.

Durka, P. J., Ircha, D., Blinowska, K. J., March 2001. Stochastic time-frequency dictionaries for matching pursuit. IEEE Tran Signal Process 49 (3), 507–510.

Ferri, R., Ferri, P., Colognola, R., Petrella, M., Musumeci, S., Bergonzi, P., 1989. Comparison between the results of an automatic and a visual scoring of sleep eeg recordings. Sleep 12 (4), 354–62.

Hae-Jeong, P., Jung-Su, O., Do-Un, J., Kwang-Suk, P., 2000. Automated sleep stage scoring using hybrid rule- and case-based reasoning. Computers and Biomedical Research 33, 330–349.

Hashizume, Y., Kuwahara, H., Uchimura, N., Mukai, M., Shirakawa, S., Satomura, T., Takeuchi, N., Tanaka, J., Maeda, H., 2001. Examination of accuracy of sleep stages by means of an automatic sleep analysis system ,,sleep ukiha". Psychiatry & Clinical Neurosciences 55 (3), 199–200.

Himanen, S.-L., Hasan, J., 2000. Limitations of Rechtschaffen and Kales. Sleep Medicine Reviews 4 (2), 149–167.

Kim, Y., Kurachi, M., et al, M. H., 1992. Agreement in visual scoring of sleep stages among laboratories in japan. Journal of Sleep Research 1, 58–60.

Larsen, L., Moe, K., Vitiello, M., Prinz, P., 1995. A note on the night-to-night stability of stages 3+4 sleep in healthy older adults: a comparison of visual and spectral evaluations of stages 3+4 sleep. Sleep 18, 7–10.

Mallat, S., 1999. A wavelet tour of signal processing, 2nd Edition. Academic Press, New York.

Mallat, S., Zhang, Z., Dec 1993. Matching pursuit with time-frequency dictionaries. IEEE Transactions on Signal Processing 41, 3397–3415.

Martin, W., Johnson, L., et al., S. V., 1972. Pattern recognition of eeg-eog as a technique for all-night sleep stage scoring. Electroencephalogr clin Neurophysiol 32 (4), 17–427.

Marzec, M., Malow, B. A., 2003. Approaches to staging sleep in polysomnographic studies with epileptic activity. Sleep Medicine 4, 409–417.

Monroe, L., 1967. Inter-rater reliability and the role of experience in scoring eeg sleep. Psychophysiol 5, 376–384.

Norman, R. G., Pal, I., Steward, C., Walsleben, J. A., Rapoport, D. M., 2000. Interobserver agreement among sleep scorers from different centers in a large dataset. Sleep 23 (7), 901–908.

Prinz, P., Larsen, L., Moe, K., Dulberg, E., Vitiello, M., 1994. C stage, automated sleep scoring: development and comparison with human sleep scoring for healthy older men and women. Sleep 17 (8), 717–7.

Rechtschaffen, A., Kales, A. (Eds.), 1968. A manual of standardized terminology, techniques and scoring system for sleep stages in human subjects. No. 204 in National Institutes of Health Publications. US Government Printing Office, Washington DC.

Schaltenbrand, N., Lengelle, R., Toussaint, M., Luthringer, R., Carelli, G., Jacqmin, A., Lainey, E., Muzet, A., Macher, J., 1996. Sleep stage scoring using the neural network model: comparison between visual and automatic analysis in normal subjects and patients. Sleep 9 (1), 26–35.

Schimicek, P., Zeitlhofer, J., Anderer, B., Saletu, P., 1994. Automatic sleep-spindle detection procedure: aspects of reliability and validity. Clinical Neurophysiology 25 (1), 26–29.

Żygierewicz, J., Blinowska, K. J., Durka, P. J., Szelenberger, W., Niemcewicz, S., Androsiuk, W., 1999. High resolution study of sleep spindles. Clin Neurophysiol 110 (12), 2136–2147.