Towards automatic sleep staging via 
Cross-Recurrence Rate of EEG and ECG activity

Nicoletta Nicolaou1,2,* Julius Georgiou1,2
1 KIOS Research Centre, University of Cyprus, Nicosia, Cyprus
2 Dept. of Electrical and Computer Engineering, University of Cyprus

*nicolett.a@ucy.ac.cy

Abstract—This paper investigates the non-linear dynamic relationship between electroencephalogram (EEG) and electrocardiogram (ECG) signals during sleep. These relationships were studied with Cross-Recurrence Rate (CRR), a non-linear measure that studies the recurrence of the phase space trajectories of dynamical systems. Data from 10 subjects during sleep were obtained from the MIT-BIH Polysomnographic database and the CRR between ECG and EEG signals was estimated. The investigations revealed strong coupling relationships between ECG and EEG that varied according to the underlying sleep stage. From a physiological perspective, the findings indicate an increase in EEG and ECG during deep sleep, while also indicating the feasibility of potential CRR application for automatic sleep staging.

Index Terms—EEG, ECG, sleep staging, cross-recurrence rate, non-linear dynamics.

I. INTRODUCTION

Human sleep can be divided into two main states - rapid eye movement (REM) and non-rapid eye movement (NREM) sleep. The latter is usually subdivided into Stages 1 and 2 (lighter sleep) and Stages 3 and 4 (deep slow wave sleep). The study of sleep and sleep-related disorders can be achieved via all-night recordings of various physiological signals, such as the electrical brain activity (EEG), heart rate (ECG) and eye movements. Most of the information currently known about sleep stems from analysis of these signals. For example, it has been found that the different sleep stages are characterised by distinct EEG patterns. The sleep staging (hypnogram) can be obtained from 20-s or 30-s EEG segments based on predefined scoring guidelines (Rechtschaffen and Kales - R&K rules [1]). The hypnogram can then be used to study the underlying sleep stages and identify any potential sleep disorders.

Sleep staging by experts is a time-consuming and somewhat subjective process and recent focus has been on automating the process using various EEG-based features, with different degrees of success [2], [3]. A general drawback of both expert scoring and automated methods is that the sleep stage scoring rules are based on changes in large segments of EEG activity only. However, the use of other physiological signals, such as heart activity and respiration, could provide additional information for a better discrimination between sleep stages [4]. This is supported by a number of studies showing that other physiological signals also vary as a function of sleep stage (e.g. variations in heart rate reported by Bonnett [5]).

In this work we investigate the feasibility of sleep staging via an advanced non-linear coupling technique that captures the intricate changes in relationships between various physiological signals during different sleep stages. More specifically, we investigate how the non-linear relationships between EEG and ECG vary during different sleep stages using cross-recurrence rate (CRR). Certain advantages of recurrence methods make them appealing over other non-linear methods: they are suitable for short time-series [6], robust to the presence of artifacts [7] and invariant to arbitrary transformations of the amplitude [7]. In addition, no assumptions on stationarity or linearity are necessary [8]. Even though recurrence methods have been applied for EEG-based sleep staging (with encouraging results) [9], [10] and exploration of ECG activity characteristics in general [11], they are yet to be applied to the study of EEG-ECG non-linear relationships during sleep.

II. METHODS

A. Dataset

The data were obtained from the MIT-BIH Polysomnographic Database [12], which is available online as part of the Physiobank archive [13]. In this study, a subset of data from 10 male subjects was used. The subjects had a mean age 41.6 (standard deviation 8.1) and were diagnosed with sleep apnoea. A total of 52h 30mins of sleep were analysed (min: 3h40', max: 6h20'). All EEG records were obtained from C3-O1. The data are sampled at 250 Hz. Each record includes an annotation file containing the expert sleep staging score, which was obtained from 30s segments using the R&K rules. These scores can be used to construct the corresponding hypnogram composed of the following 6 stages: sleep stages 1-4, REM sleep, and wakefulness. Figures 1 and 2 show example short segments of EEG and ECG signals from two randomly chosen subjects.

B. Cross-Recurrence Rate

The CRR allows us to study the recurrent dynamics between different systems by estimating how many times a particular state occurs simultaneously in both systems, at a particular time delay. It can be considered as the bivariate equivalent of the recurrence rate [14] and is, thus, estimated in an
analogous manner. To estimate the CRR the data is converted to a symbolic sequence and encoded into order patterns. The symbolic sequence is obtained from time-delay embedding of the signals, which effectively corresponds to a reconstruction of the original phase-space \cite{15}. Therefore, for a scalar time series,

\[
X = [x(1), x(2), \ldots, x(T)],
\]

\[
\hat{x}(t) = (x(t), x(t-\tau_{ed}), \ldots, x(t-(m-1)\tau_{ed}))
\]

where \(\hat{x}(\cdot)\) is the embedded time series, and \(m\) and \(\tau_{ed}\) are the embedding dimension and embedding delay respectively. A rank is then assigned to the resulting \(m\)-dimensional EEG amplitudes from a possible \(m!\) order patterns, \(\pi_{k}^x(t), k = 1, \ldots, m!\) (for example, Table I shows the \(m!\) patterns for \(m = 3\)). The recurrence of the phase-space trajectories of two signals, \(\hat{x}_1(t), \hat{x}_2(t)\), to the same order patterns within a predefined delay, \(\tau_{ord}\), is described by the recurrence matrix,

\[
R(t_j, t_j + \tau_{ord}) = \begin{cases} 1, & \pi^{\hat{x}_1}_{k_1}(t_j) = \pi^{\hat{x}_2}_{k_2}(t_j + \tau_{ord}) \\ 0, & \text{otherwise} \end{cases}
\]

where \(j = 1, \ldots, N; N = T - m - \tau_{ord} + 2\). The CRR is, then, estimated as follows:

\[
CRR(\hat{x}_1(t), \hat{x}_2(t + \tau_{ord})) = \frac{1}{N} \sum_{j=1}^{N} R(t_j, t_j + \tau_{ord})
\]

The factor \(N\) is a normalizing factor such that \(0 \leq CRR \leq 1\). The choice of parameters \(m\) and \(\tau_{ord}\) is important. In these preliminary investigations we chose the values \(m = 3\) and \(\tau_{ord} = 1\), which were shown to be appropriate for EEG signals (e.g. see \cite{14}).

### Table I

<table>
<thead>
<tr>
<th>(x(t))</th>
<th>(k = 1)</th>
<th>(k = 2)</th>
<th>(k = 3)</th>
<th>(k = 4)</th>
<th>(k = 5)</th>
<th>(k = 6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\pi_{k}^x(t))</td>
<td>012</td>
<td>021</td>
<td>021</td>
<td>021</td>
<td>021</td>
<td>102</td>
</tr>
</tbody>
</table>

#### C. Methodology

The CRR is estimated from non-overlapping 5s segments. In addition to the polysomnograms, each record is accompanied by corresponding sleep stage annotations obtained from 30s segments based on the R&K sleep scoring rules. Using the annotations provided for each subject, the mean CRR (\(CRR_{\mu}\)) and standard deviation (\(CRR_{\sigma}\)) values representing each of the 5 sleep stages, \(i = 1, \ldots, 5\), and wakefulness, \(i = 6\), are obtained. A subject-wise mean CRR (with standard deviation) during each sleep stage can then be estimated.

#### III. RESULTS AND DISCUSSION

Figures 1 and 2 show example short segments of (a) EEG and (b) ECG signals from two randomly chosen subjects (slp45 and slp48 respectively), together with the (c) hypnogram constructed from the corresponding annotations. In figures 1(d) and 2(d) the estimated CRR values are overlaid on the hypnogram (the CRR values are scaled and smoothed with a moving average filter for visualisation purposes). The CRR values exhibit distinct changes during different sleep stages and capture sleep microstructure, something which is not possible with R&K-based scoring.
Figure 3 shows the grand CRR mean over all subjects (bars) plotted with ± one standard deviation. The large standard deviation is a reflection of the large inter-subject variability in the actual CRR values. Overall, wakefulness seems to display the lowest CRR values, which increase progressively as the subject falls asleep. The difference in CRR values during REM sleep and wakefulness could be an indication that, despite the similar EEG activity in both states, the two are distinct processes characterised by different relationships between EEG and ECG activity. It is important to consider that the mean CRR values presented in Table II and fig. 3 were estimated based on the sleep staging scores resulting from the ReK rules. However, the CRR is estimated on much shorter data segments (5s). Given that it is highly possible that there could be brief fluctuations of different nature between different sleep stages within each 30s segment for each patient, it is more accurate to consider the individual CRR values (as shown in figs. 1(d) and 2(d)) rather than the CRR mean for a particular sleep stage.

While there are studies reporting isolated changes in EEG and ECG activity during sleep, there are no studies of their non-linear co-variation. The use of CRR allows the study

<table>
<thead>
<tr>
<th>Subj.</th>
<th>W</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>REM</th>
</tr>
</thead>
<tbody>
<tr>
<td>slp03</td>
<td>0.197</td>
<td>0.209</td>
<td>0.220</td>
<td>0.231</td>
<td>—</td>
<td>0.218</td>
</tr>
<tr>
<td>slp04</td>
<td>0.211</td>
<td>0.228</td>
<td>0.254</td>
<td>0.272</td>
<td>—</td>
<td>0.267</td>
</tr>
<tr>
<td>slp14</td>
<td>0.199</td>
<td>0.204</td>
<td>0.203</td>
<td>0.202</td>
<td>0.191</td>
<td>0.184</td>
</tr>
<tr>
<td>slp16</td>
<td>0.192</td>
<td>0.199</td>
<td>0.200</td>
<td>0.205</td>
<td>0.206</td>
<td>0.202</td>
</tr>
<tr>
<td>slp45</td>
<td>0.194</td>
<td>0.203</td>
<td>0.221</td>
<td>0.219</td>
<td>0.244</td>
<td>0.220</td>
</tr>
<tr>
<td>slp48</td>
<td>0.184</td>
<td>0.201</td>
<td>0.230</td>
<td>0.306</td>
<td>—</td>
<td>0.309</td>
</tr>
<tr>
<td>slp59</td>
<td>0.146</td>
<td>0.145</td>
<td>0.140</td>
<td>0.134</td>
<td>0.134</td>
<td>0.138</td>
</tr>
<tr>
<td>slp60</td>
<td>0.171</td>
<td>0.169</td>
<td>0.171</td>
<td>—</td>
<td>—</td>
<td>0.176</td>
</tr>
<tr>
<td>slp61</td>
<td>0.149</td>
<td>0.146</td>
<td>0.143</td>
<td>0.133</td>
<td>—</td>
<td>0.137</td>
</tr>
<tr>
<td>slp66</td>
<td>0.178</td>
<td>0.172</td>
<td>0.172</td>
<td>0.180</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>TOTAL</td>
<td>0.182</td>
<td>0.189</td>
<td>0.196</td>
<td>0.210</td>
<td>0.194</td>
<td>0.206</td>
</tr>
</tbody>
</table>
of similarities in the phase-space trajectories of both signals together: lower CRR values indicate less similarity of the phase-space trajectories of the brain and heart activity. Lower values are seen in lighter sleep stages (stages 1 and 2). The increase of CRR values during deeper sleep stages indicates an increase in the similarity of the studied activity. These observations tie in with findings from other studies, where it has been shown that EEG complexity decreases as the sleep deepens and increases again during REM sleep [10], while ECG activity exhibits a similar decrease in complexity during deep sleep compared to REM sleep [16]. Our observations could also be related to changes in the frequency content of both signals during different sleep stages [17].

The proposed methodology has a number of advantages: (1) In the R&K scoring system, the designated sleep stage is that which is present for over half of the 30-s segment, therefore it characterizes the dominant sleep stage while failing to capture the dynamic structure of sleep. The proposed methodology allows the study of short segments and captures sleep microstructure (see figures 1 and 2) in a way that the R&K rules cannot. Taking into account the sleep microstructure can lead to more accurate hypnograms and sleep staging; (2) The co-variation of two types of physiological activity during sleep is taken into account. Given that different stages of sleep cause various changes to physiological activity, more information is being utilized and a more accurate sleep staging score can be obtained; (3) By modifying the time delay between the studied signals we can also investigate the lead-lag relationships. This is possible due to the bidirectional nature of the CRR. This is an important advantage over other methods that are not bidirectional, as using lead-lag information could lead to more accurate sleep staging.

The particular CRR patterns were estimated from C3-O1 electrode location. However, it is important that the proposed methodology is validated for different spatial locations, as the changes during different sleep stages are more prominent in different areas of the brain. In addition, the choice of parameters \( m \) and \( \tau_{\text{ord}} \) is important. In these preliminary investigations we chose values that were shown to be appropriate for EEG applications. It is possible that other combinations of these parameters may reveal different aspects of the relationship between EEG and ECG signals. An additional consideration is the presence of artifacts in the signals, since no artifact removal has been performed here. Even though recurrence methods are robust to artifacts, some outlier CRR values could be expected. Therefore, the assessment of the effect of artifacts to CRR robustness is something that should also be considered. In addition, the particular database utilized contains data only from male subjects and it would be interesting to investigate whether significant gender-related differences exist. Finally, the mean CRR values during S4 sleep should be interpreted with caution, as only 4 subjects exhibited S4 sleep.

**IV. CONCLUSION**

The application of cross-recurrence rate between EEG and ECG activity as a means of sleep staging is investigated. The CRR values show clear changes during different sleep stages, supporting the feasibility of using a CRR-based method for automatic sleep staging. The development and validation of such a measure will be part of future work. Furthermore, identifiable changes in coupling between additional physiological signals during sleep, e.g. muscle activity, eye movements and respiration, will be investigated. This will allow a more general picture of what changes occur during the different sleep stages, with the aim of using this information for the development of a more accurate automatic sleep staging methodology.

**REFERENCES**


