The use of wearable sensors has been mainly limited to coarse-grained methods such as actigraphy, for which limb motions are logged providing some insight in patient's sleep. Actigraphy correlates well with PSG in differentiating sleep from wake [103], but rather than a replacement for PSG, it should be regarded as another means for assessing sleep, particularly when sleep architecture and extensive physiological monitoring are not necessary.

Activity monitors are useful tools that are becoming popular to objectively assess the sleep–wake cycle. They provide minimally invasive measures of the continuity and hence quality of sleep and other physiological measurements such as energy expenditure, body temperature and galvanic skin response. In this work, metabolic and physiological data recorded during night-time using an activity monitor are symbolized and presented as the "letters" composing the "words" that describe the night of a subject. The co-occurrence of these words in different ways and proportions during the night creates groups of words describing the modalities in which a subject sleeps. Using these data we extracted patients' sleep modalities that were valid to assess sleep in relation to the presence of the pathological condition. While previous studies aimed at finding differences between healthy individuals and patients with COPD [98], this study seeks to evaluate severity classification of the disease. This is a challenge because differences in objective sleep measures between COPD classes are subtle [98, 99] making them difficult to detect. The proposed methodology, exclusively defined using data coming from one unobtrusive device, is able to differentiate between COPD and healthy-type of nights, and to discriminate between different GOLD grades and MMRC scores.

7.3 Background

Topic models are algorithms for discovering the hidden grouping variables that pervade a large and unstructured collection of documents. LDA is an example of a topic model [139] in which data are treated as observations arising from a generative probabilistic process. In the context of text modelling, given a set of topics defined as distributions over words, the generative process populates the documents with words such that the documents have a particular desired thematic structure. Beside its generative process, LDA can also be used to calculate the hidden variables that likely generated the collection of documents. One of the ways to achieve this is to use variational inference to approximate the posterior distribution over the hidden variables defined by LDA. In a nutshell, variational inference posits a parametrized family of distributions over the hidden structure, and then, finds the member of that family that is closest to the posterior according to the *Kullback–Leibler* divergence. The intuition behind using LDA [139] for sleep monitoring is that each night is a mixture of thematically coherent measures just as a text document is a mixture of thematically coherent words. The graphical model for LDA is shown in Figure 43. All the assessed nights ($d_{1:D}$) share the same set of sleep modalities ($\beta_{1:K}$) that are defined as Dirichlet distributions over the observed set of symbols (W) which are the terms of a fixed vocabulary. The observed symbols (input of the model) are composed by multimodal measures coming from the sensors of an activity monitor. Each assessed night exhibits sleep modalities in different proportion providing an explicit finger print ϑ. In particular, each night is a different distribution ($\vartheta_{1:D}$) over the sleep modalities activation probabilities that also follows a Dirichlet distribution. In such a model, the N symbols ($W_{d,n_{1:N}}$) that compose the D nights are the only random variables observed and depend on the per word sleep modality assignment ($Z_{d,n}$) and all the β_k . Each sleep modality then is composed indirectly by low-level