# watson: An R Package for Fitting Mixtures of Watson Distributions

Lukas Sablica[1], Kurt Hornik[1], and Josef Leydold[1]

[1]Institute for Statistics and Mathematics, WU Wirtschaftsuniversität Wien, 1020 Vienna, Austria

In our work we present and showcase the R package **watson** which provides a computational framework for fitting and random sampling of the Watson distribution on a $p$-dimensional sphere. We first introduce the random sampling scheme of the package, which offers two sampling algorithms that are based of the results of [1]. What is more, the package offers a smart tool to combine these two methods, and based on the selected parameters, it approximates the relative sampling speed for both methods and picks the faster one.

In addition, we describe the main fitting function for the mixtures of Watson distribution which uses the expectation-maximization (EM) algorithm. Special features are the possibility to use multiple variants of the E-step and M-step, sparse matrices for the data representation and a control parameter which will dynamically eliminate small clusters with overall contribution smaller than this parameter. Moreover, we discuss the numerical issues of the whole fitting procedure and describe how this is handled and solved in the package. Finally, we demonstrate the package on multiple examples involving misspecified simulation study, estimation of the New Zealand earthquake data and depth image clustering.

[1] L. Sablica, K. Hornik, and J. Leydold, "Random sampling from the watson distribution in arbitrary dimensions," *Under Review in: Journal of Multivariate Analysis*, 2023.