TENSOR-BASED BLIND SYSTEM IDENTIFICATION WITH FOCUS ON BIG DATA

*Martijn Boussé**

*Department of Electrical Engineering (ESAT),
KU Leuven, Kasteelpark Arenberg 10, 3001 Leuven, Belgium.

ABSTRACT

In blind system identification (BSI) one tries to identify an unknown system based only on the observed outputs. This problem arises in many applications because one often does not know the inputs or because they are difficult to measure. Recent advances in several applications within fields such as biomedical sciences and sensor array processing have lead to a large increase in both the number of samples and sensors. Conventional BSI methods are not applicable in such a big data setting because they depend strongly, in some cases even exponentially, on the number of sensors. In this doctoral research, we will develop new methods to blindly identify large-scale systems based on tensors which are a natural generalization of vectors and matrices. Specifically, we use tensor decompositions as low-rank parametric representations for the system coefficients, enabling us to model the system in a very compact way. A similar strategy has proven successful in scientific computing, allowing to solve problems in a number of unknowns that exceeds the number of atoms in the universe.

Blind system identification (BSI), illustrated in Figure 1, is an important problem with a lot of applications in fields such as biomedical sciences, wireless communications, audio and image processing, sensor array processing, and seismic prospection. Focusing specifically on biomedical applications, we see a large increase in the number of samples and sensors. Consider for example electroencephalography (EEG) caps which are promising tools for long-term neuromonitoring; such systems can already have up to 250 sensors. Other examples include electrocorticography (ECoG) with high spatial resolution, surface electromyography (sEMG), and neural dust with thousands of miniature sensors dispersed throughout the outer layer of the brain. The important conclusion here is that we see a large increase of data in both the spatial and temporal domain.

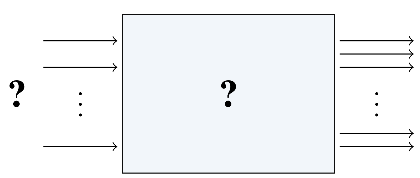

Fig. 1. Visualization of blind system identification.

Fig. 2. Visualization of a vector (left), matrix (middle), and third-order tensor (right).

Higher-order tensors are natural generalizations of vectors (first order) and matrices (second order), see Figure 2. Tensor-based analysis has been strongly rising in popularity in recent years because tensor tools are much more powerful than conventional linear techniques but far more interpretable than nonlinear methods. We know that matrices can be decomposed in a sum of *intrinsically small* terms or *atoms*. If such a matrix has a low-rank structure only a few of such atoms are needed in the sum, enabling a very compact representation. Interestingly, one can do something similar for tensors but now the representation becomes much more compact with increasing order, indicating the power of tensors for big data. This idea is not new but has been applied in tensor-based scientific computing in high dimensions, where one is able to work with functions that have more unknown values that atoms in the universe. The novelty is that we apply this idea to handle large-scale problems in a novel way for BSI.

It is clear that tensors offer some remarkable possibilities to combine BSI and big data. The strategy is as follows. First, we *fold* the observed data matrix (i.e., the measured outputs) into a tensor by a process called tensorization. Second, we decompose the resulting tensor in a sum of only a few atoms. This is possible because the low-rank structure is often present in big data systems. It can then be shown that the parameters of the model that we use to identify the system as well as the inputs can be derived from these atoms. This provides a highly compact representation of the system because of the small amount of (intrinsically small) atoms. This strategy is completely new and different from known methods. We apply it to several kinds of well-known models, leading to new BSI methods that are applicable in a big data setting in contrast with conventional techniques. Our methods are highly generic and applicable in many applications.

978-0-9928626-3-3/15/$31.00 ©2015 IEEE