

The Research Group

**Digital Mathematics**

has the honor to invite you to the public defense of the PhD thesis of

## **Gilles Inghelbrecht**

to obtain the degree of Doctor of Sciences

Title of the PhD thesis:  
**Parallelization of Least Squares Systems of Equations under the Matrix Splitting Theory**

Promotors:

**Prof. dr. Kurt Barbé (VUB)**

The defense will take place on  
**Monday, June 27, 2022 at 16h in  
auditorium D.0.05**

**Members of the jury**

Prof. dr. Philippe Cara (VUB, chair)  
Prof. dr. Jan De Beule (VUB, secretary)  
Prof. dr. Ann Nowé (VUB)  
Prof. dr. Bart De Moor (KU Leuven)  
Prof. dr. Paolo Carbone (Perugia University)  
Prof. dr. Rosemary Renaut (Arizona State University)

### **Curriculum vitae**

**Gilles Inghelbrecht** was born in Ukkel, Belgium, in 1992. He received the M. Sc degree in Mathematics at the University of Antwerp in 2015. He joined the research group Digital Mathematics (DIMA) at the Vrije Universiteit Brussel as a PhD student. He was a recipient of the Graduate Fellowship Award from the IEEE Instrumentation and Measurement Society in 2017.

### **Abstract of the PhD research**

In our society today, we are faced with an immeasurable amount of digital information. The analysis of this data provides insights, both on an economic and on a scientific level. The large influx of data helps to identify trends or to deliver new scientific results. The resulting datasets, however, are potentially too large, such that the required statistical analysis becomes impossible to be solved numerically in a timely fashion.

In this thesis we focus on regression analysis, which constitutes a standard method within statistics and machine learning. In a big data context, both the number of observations and the number of variables in these models may increase without limit. A numerically challenging problem within the theory of regression analysis is the solution of the least squares estimator.

During this project, we focussed on the least squares multisplitting (LSMS) method, which allows a large regression problem to be replaced by a set of smaller regression problems, where the splitting happens within the variables. These smaller regression problems are solvable in parallel and communicate with each other to update their local solutions in an iterative fashion. An important shortcoming for an end-user is the lack of clarity in how this variable splitting should be performed. In order to address this partition problem, we applied clustering and graph-theory to provide a solution. We further developed alternative hierarchical methods to iteratively solve the splitting-problem and studied the convergence behaviour.

An important application domain for our methods is signal processing. More precisely, we considered a least squares problem that arises in the estimation of the frequency response function under missing samples. This type of problem has a beautiful statistical solution based on local polynomial approximation. We show that an application of the LSMS provides improvement, but we further developed specific solvers using 2-cyclic matrix splittings, which optimally employ the structure of this regression problem. The resulting method improves the arithmetic complexity from cubic to (iterative) log-linear, such that the numerical scalability is no longer an issue.