
Computing Floating-Point Matrix Products via Low-Precision Integer Arithmetic

Massimiliano Fasi^{*1}, Ahmad Abdelfattah , Jack Dongarra , Mantas Mikaitis , and Françoise Tisseur

¹University of Leeds – United Kingdom

Abstract

Matrix multiplication is a fundamental kernel in scientific computing, and efficient implementations underpin the performance of many numerical algorithms for linear algebra. The Ozaki scheme is a method that computes matrix–matrix products by recasting them as sequences of error-free computations. First developed in 2008 in the context of summation, this technique has recently seen a resurgence of interest, because it is particularly well suited to the mixed-precision matrix-multiplication units available on modern hardware accelerators. Latest-generation accelerators are particularly efficient at computing products between matrices of low-precision integers. In scientific applications, integers are typically not sufficient, but, in the last couple of years, variants of the Ozaki scheme that rewrite floating-point matrix multiplications in terms of integer matrix products have been proposed. Using error analysis, we characterise the conditions in which these methods can fail, and we propose input-dependent ways to obtain accuracy-performance tradeoffs.

^{*}Speaker