

---

# How Tensor Core-Accelerated Mixed-Precision Computing and Floating-Point Emulation Can Improve Linear Algebra Performance and Energy Efficiency

Benedikt Dorschner\*<sup>1</sup>, Harun Bayraktar<sup>1</sup>, John Gunnels<sup>1</sup>, and Samuel Rodriguez Bernabeu<sup>1</sup>

<sup>1</sup>NVIDIA – United States

## Abstract

The reduced- and mixed-precision computing capabilities of GPUs have grown rapidly in recent years which has been primarily driven by AI-based applications (e.g., LLMs). These processors demonstrate an outsized power-efficiency (FLOPS/watt) advantage for matrix multiplications over systems almost exclusively focused upon native single- and double-precision arithmetic. Thus, this presents both a great opportunity and motivation to leverage these capabilities for dense linear algebra, through the use of various mixed-precision algorithms and floating-point emulation techniques, to facilitate greater scientific computing throughput without sacrificing accuracy. We'll touch upon a number of these approaches such as the Ozaki-I and Ozaki-II schemes for double-precision matrix multiplication emulation, and present real-world case studies that provide compelling evidence in support of this path to further increase the science per watt of GPU accelerated computing.

---

\*Speaker