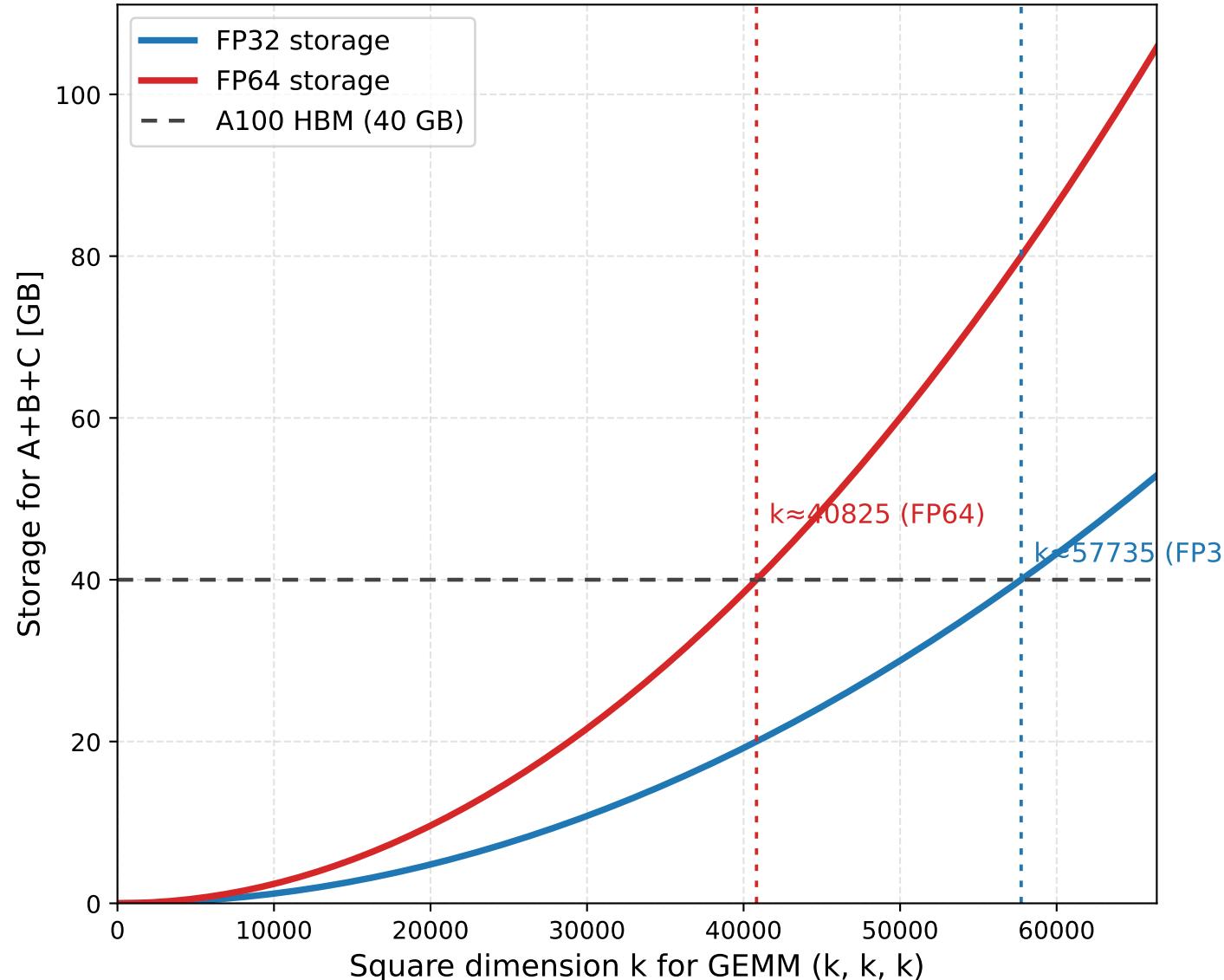


# GEMM Storage Capacity Model for A100 (A, B, C only; no workspace)

## Global-Memory Footprint vs A100 HBM



# Same Footprint Model vs Per-SM Shared-Memory Limit

