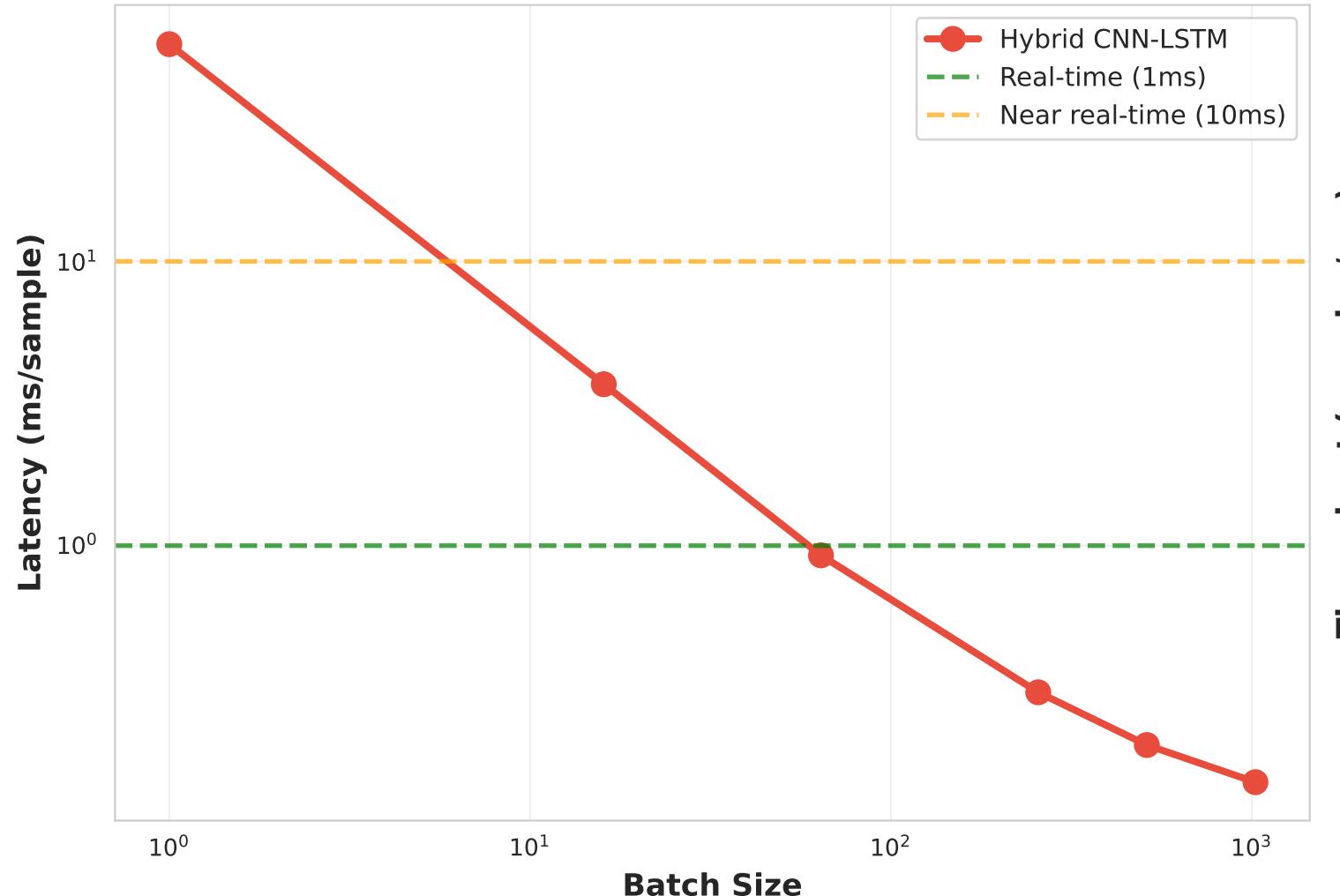


Inference Latency vs Batch Size



Inference Throughput vs Batch Size

