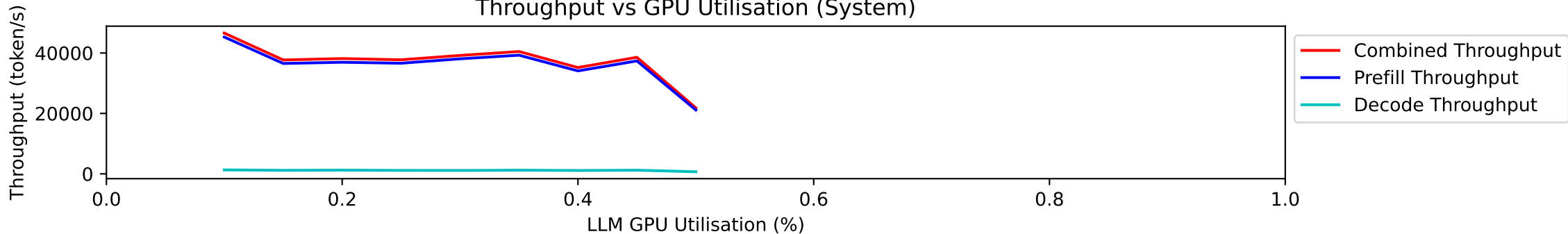
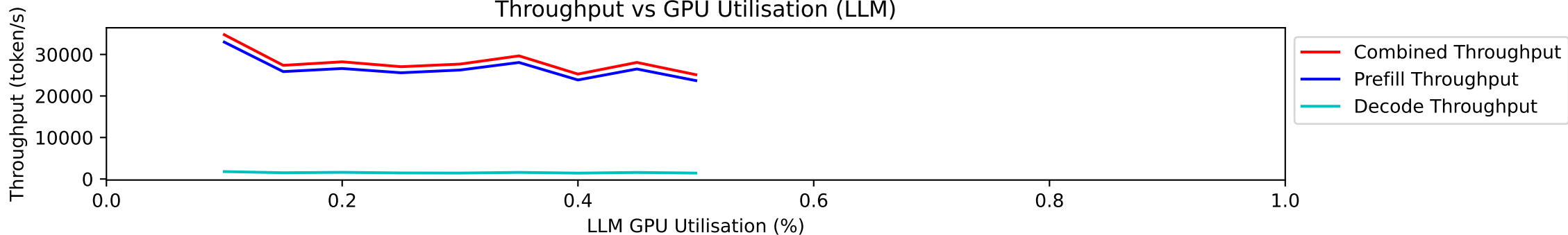


Throughput vs GPU Utilisation (System)



Throughput vs GPU Utilisation (LLM)



Throughput vs GPU Utilisation (PRM)

