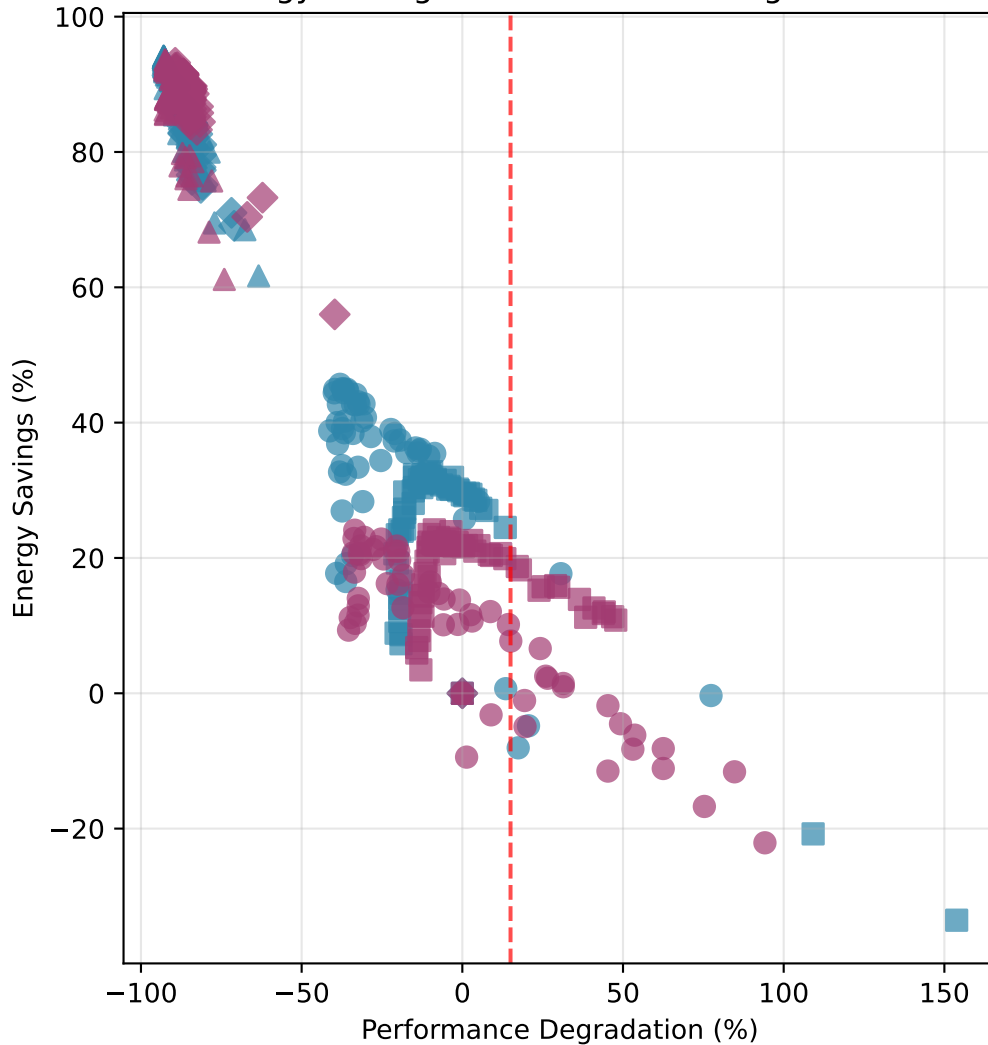


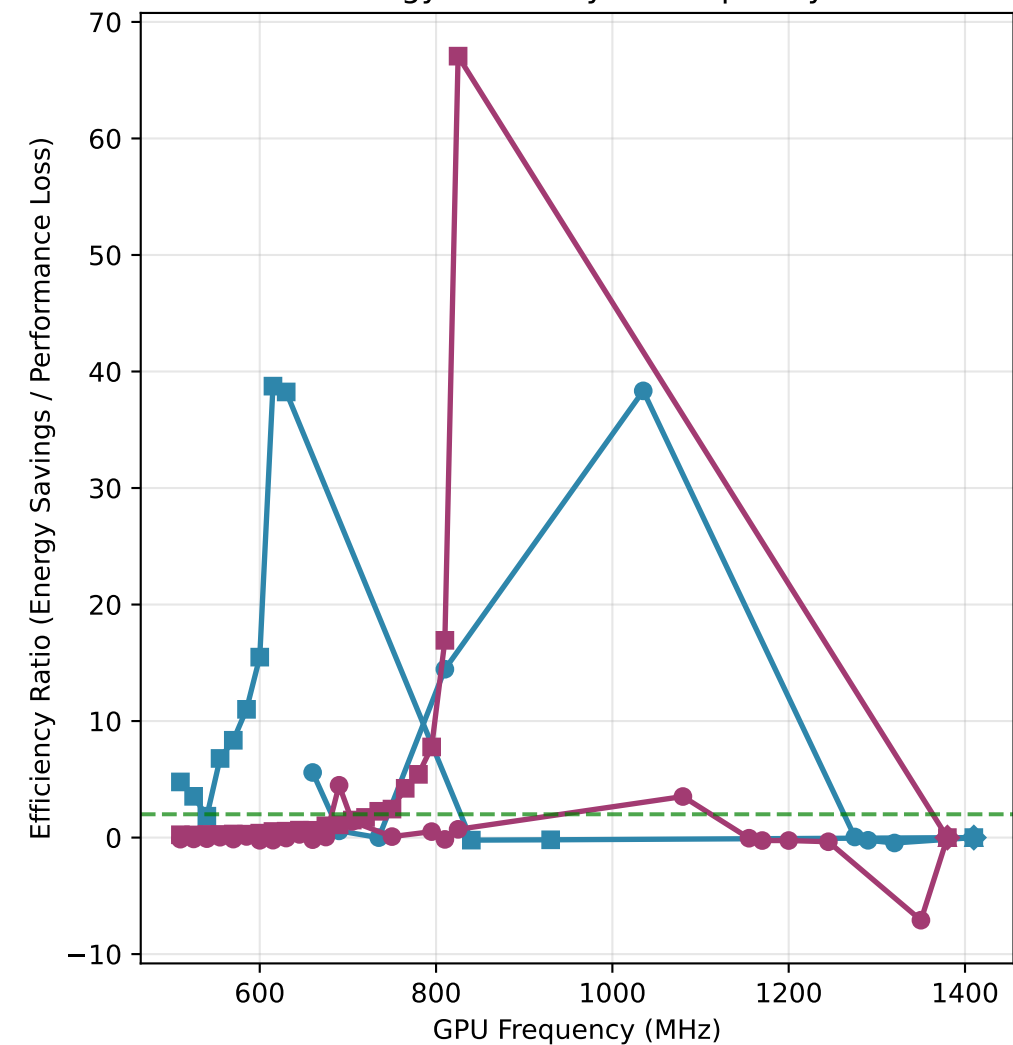
GPU Frequency Optimization: Performance vs Energy Trade-offs

Energy Savings vs Performance Degradation



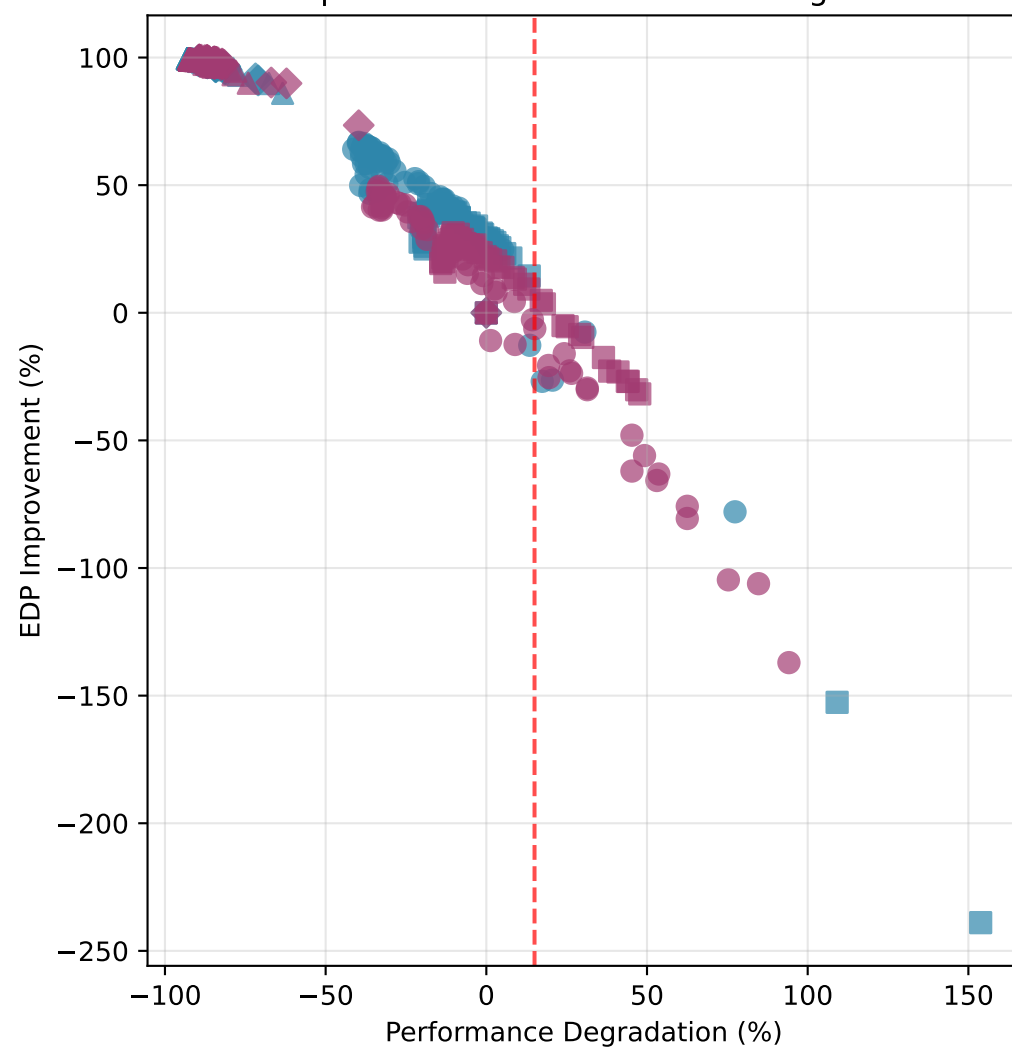
- A100+LLAMA
- A100+STABLEDIFFUSION
- A100+VIT
- A100+WHISPER
- V100+LLAMA
- V100+STABLEDIFFUSION
- V100+VIT
- V100+WHISPER
- 15% Performance Limit

Energy Efficiency vs Frequency



- A100+LLAMA
- A100+STABLEDIFFUSION
- A100+VIT
- A100+WHISPER
- V100+LLAMA
- V100+STABLEDIFFUSION
- V100+VIT
- V100+WHISPER
- 2:1 Efficiency Target

EDP Improvement vs Performance Degradation



Energy vs Power Reduction

