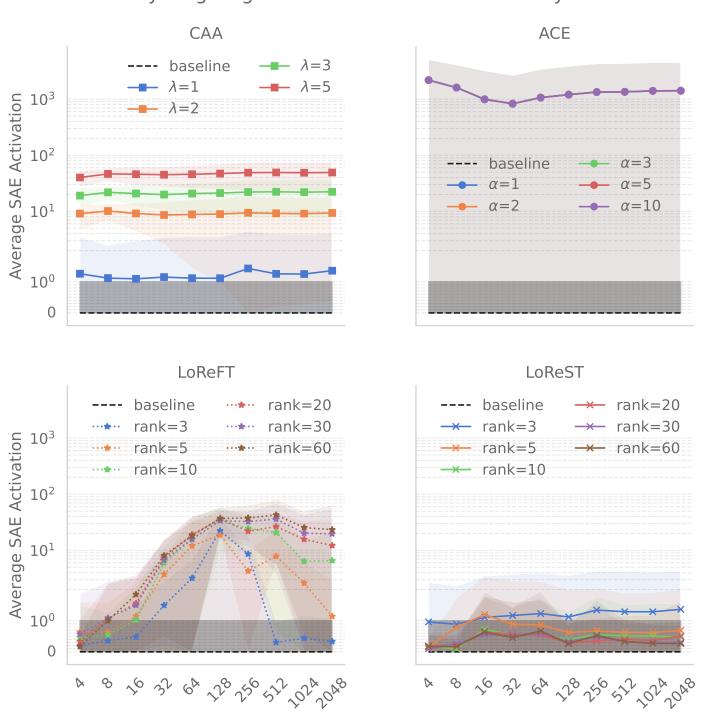
Symlog Target SAE Features for GPT-2 at Layer 7



Number of steering examples

Number of steering examples