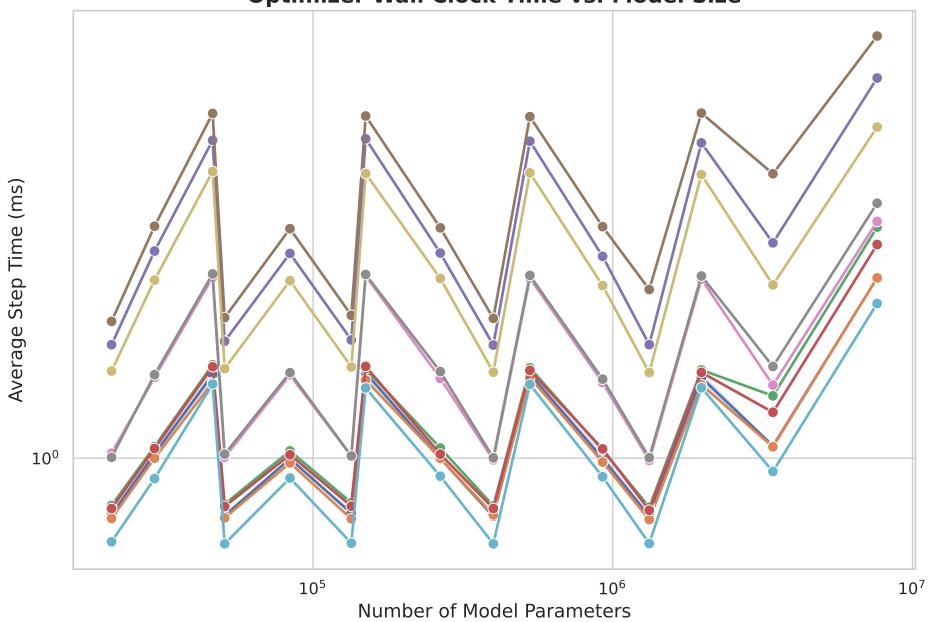
Optimizer Wall Clock Time vs. Model Size



Optimizer | Device optimizer

- --- adagrad
- --- rmsprop
- **-●-** amsgrad
- --- adam
- **--** adan
- **-●-** adamp
- **--** lion
- **-●-** madgrad
- --- diffgrad
- pulsegrad device
- **-**●**-** cuda