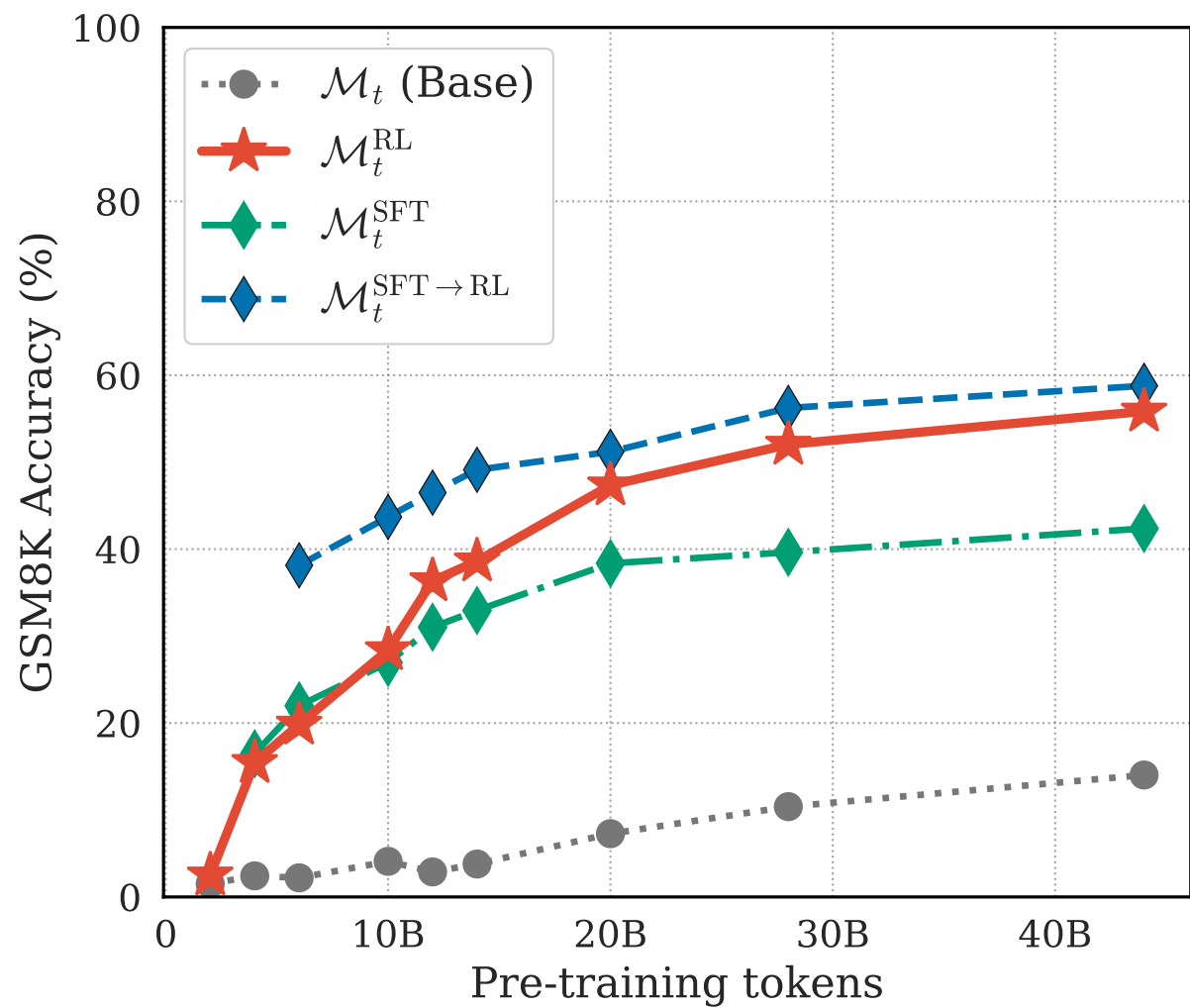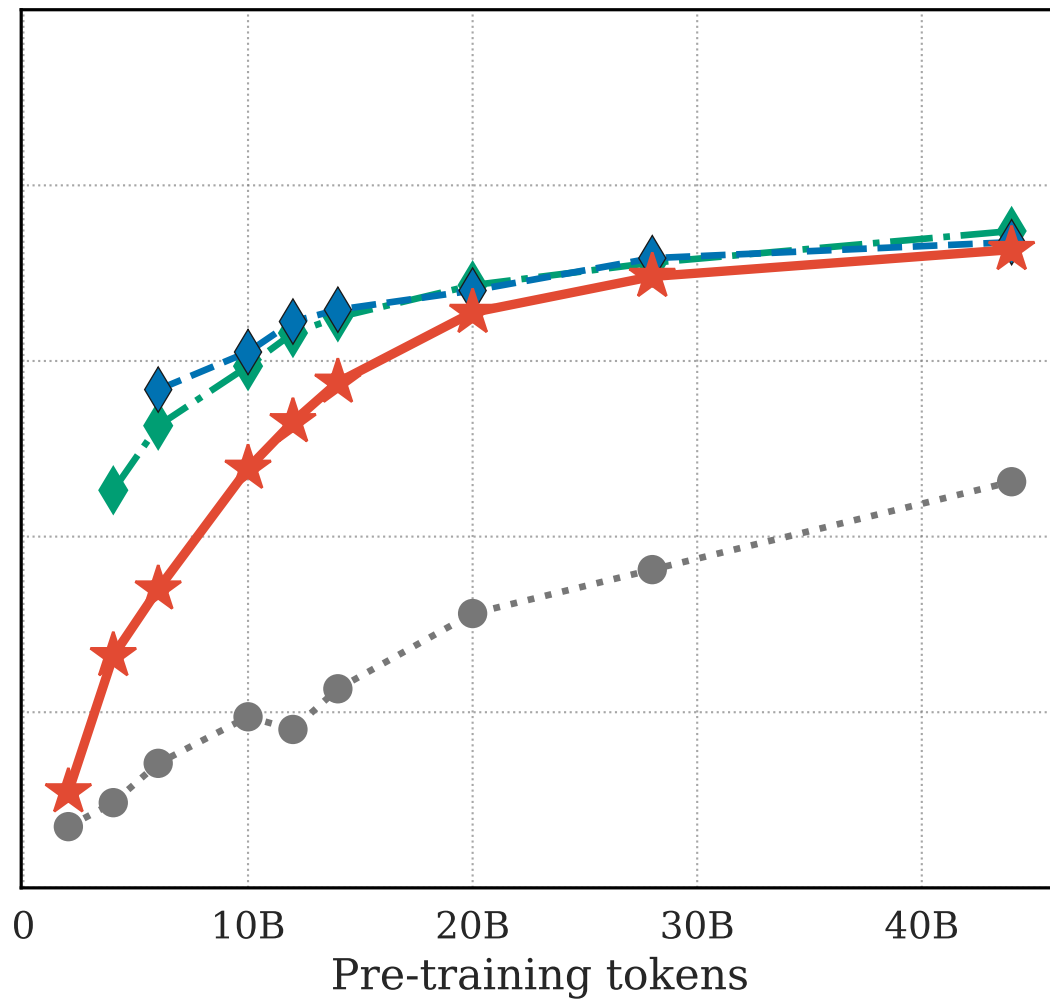**Pass@1**   **Pass@8**   **Pass@32**

Legend:
- $\mathcal{M}_t$ (Base)
- $\mathcal{M}_t^{\mathrm{RL}}$
- $\mathcal{M}_t^{\mathrm{SFT}}$
- $\mathcal{M}_t^{\mathrm{SFT} \rightarrow \mathrm{RL}}$

Axes: GSM8K Accuracy (%) vs. Pre-training tokens