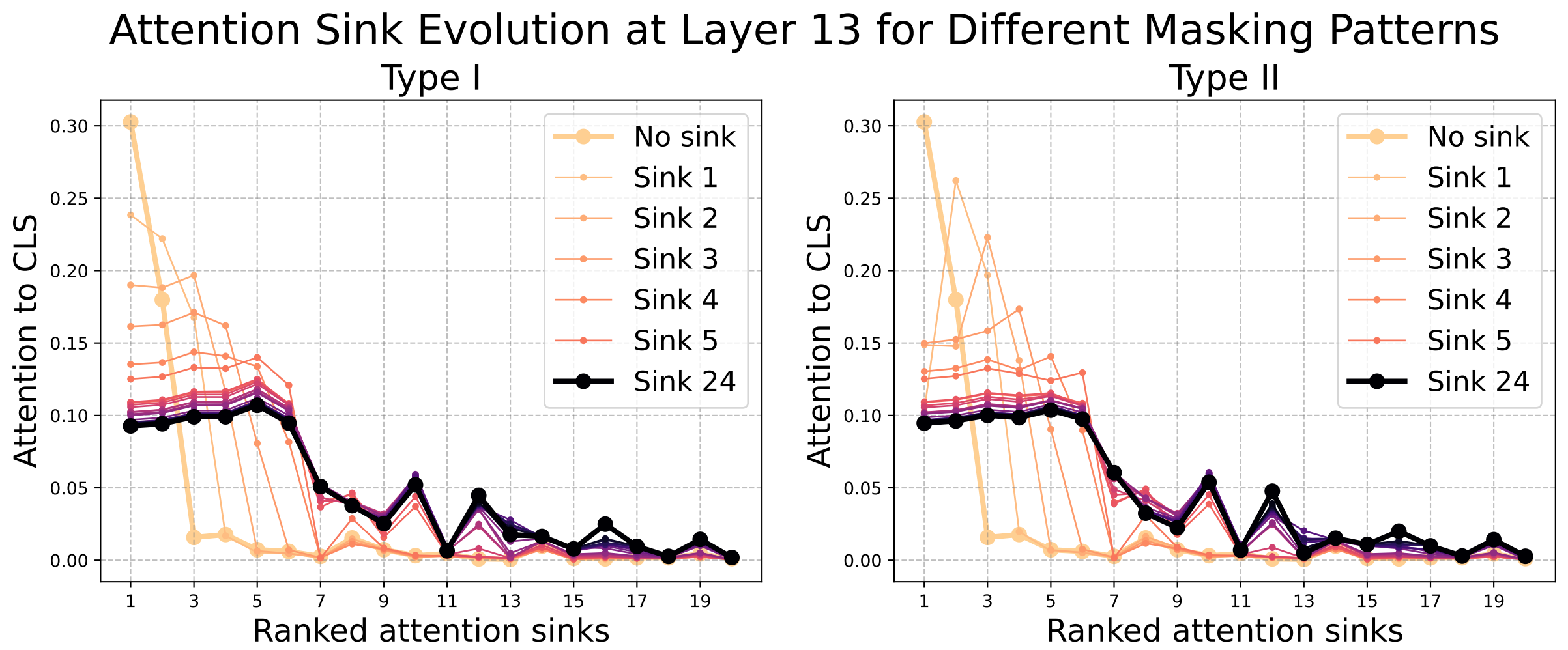
Attention Sink Evolution at Layer 13 for Different Masking Patterns
Type I
Type II
Attention to CLS
Ranked attention sinks
No sink
Sink 1
Sink 2
Sink 3
Sink 4
Sink 5
Sink 24
0.00
0.05
0.10
0.15
0.20
0.25
0.30
1
3
5
7
9
11
13
15
17
19