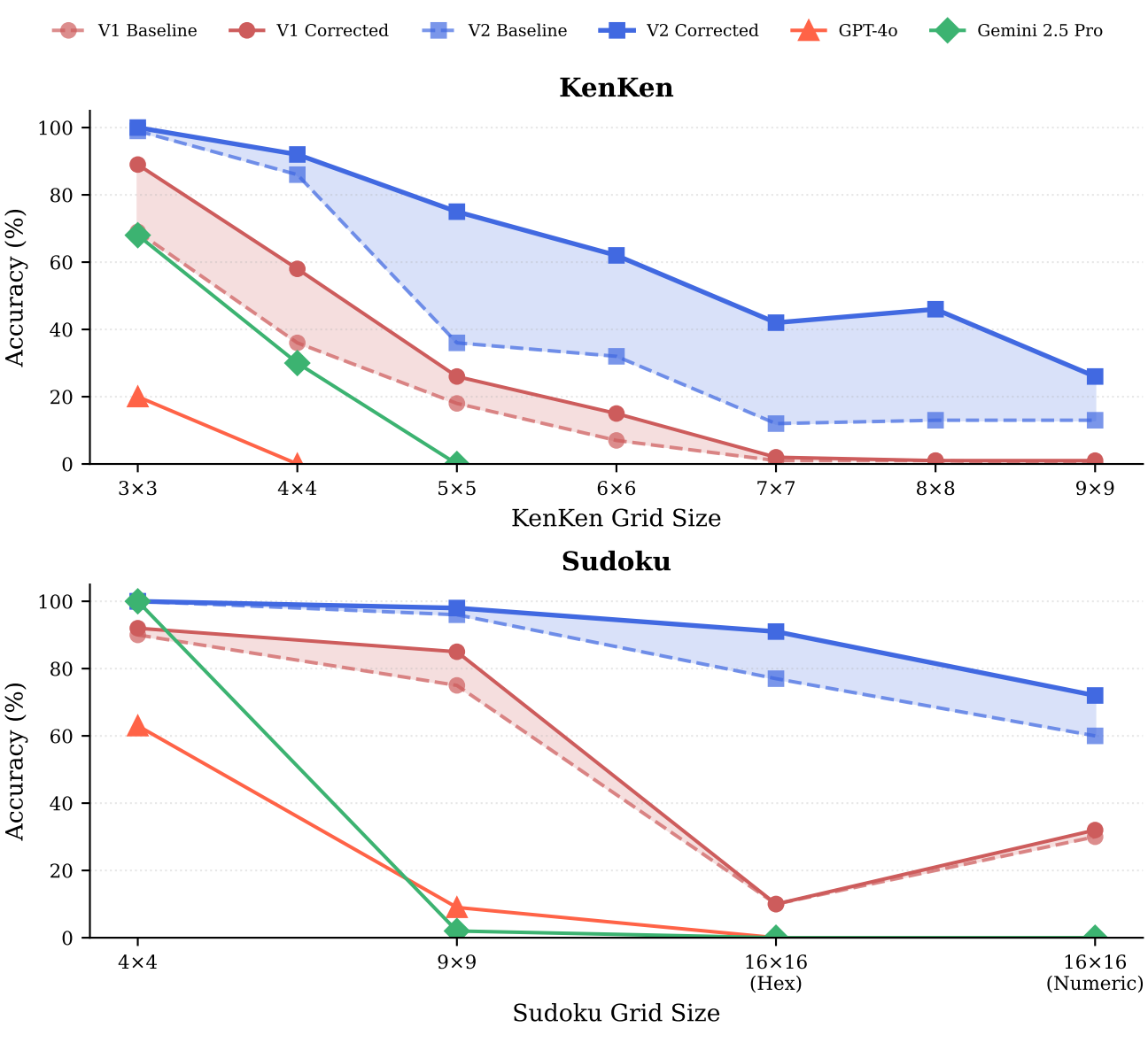
V1 Baseline
V1 Corrected
V2 Baseline
V2 Corrected
GPT-4o
Gemini 2.5 Pro
KenKen
Accuracy (%)
100
80
60
40
20
0
3×3
4×4
5×5
6×6
7×7
8×8
9×9
KenKen Grid Size
Sudoku
Accuracy (%)
100
80
60
40
20
0
4×4
9×9
16×16
(Hex)
16×16
(Numeric)
Sudoku Grid Size