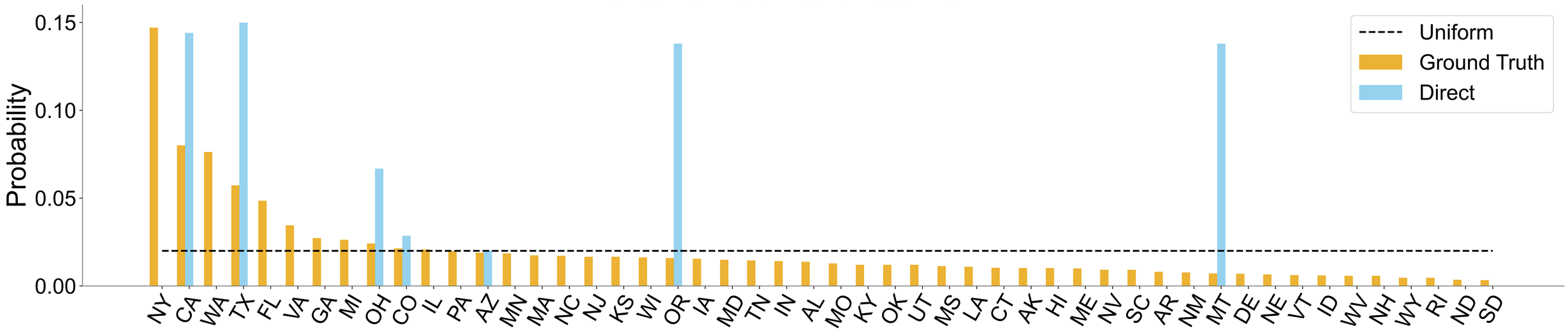
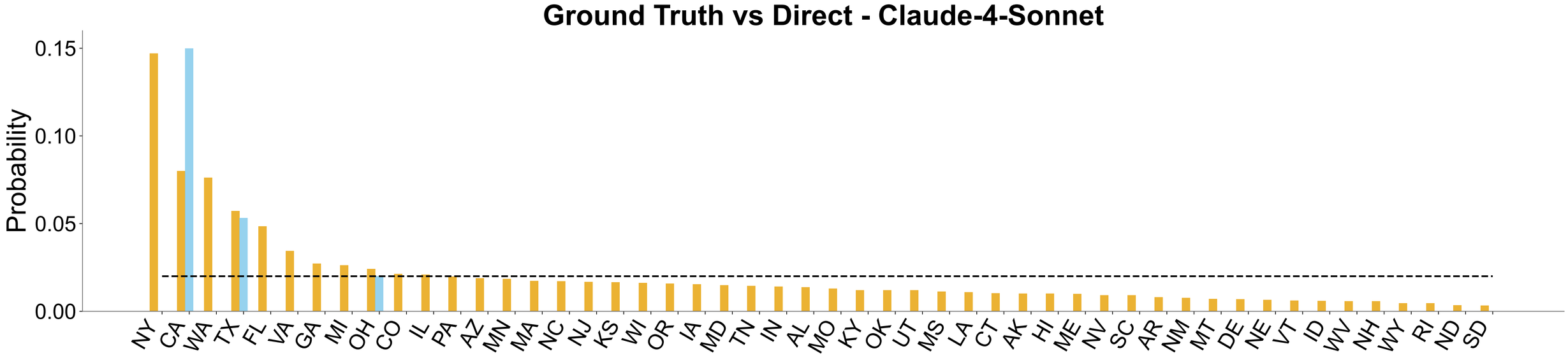
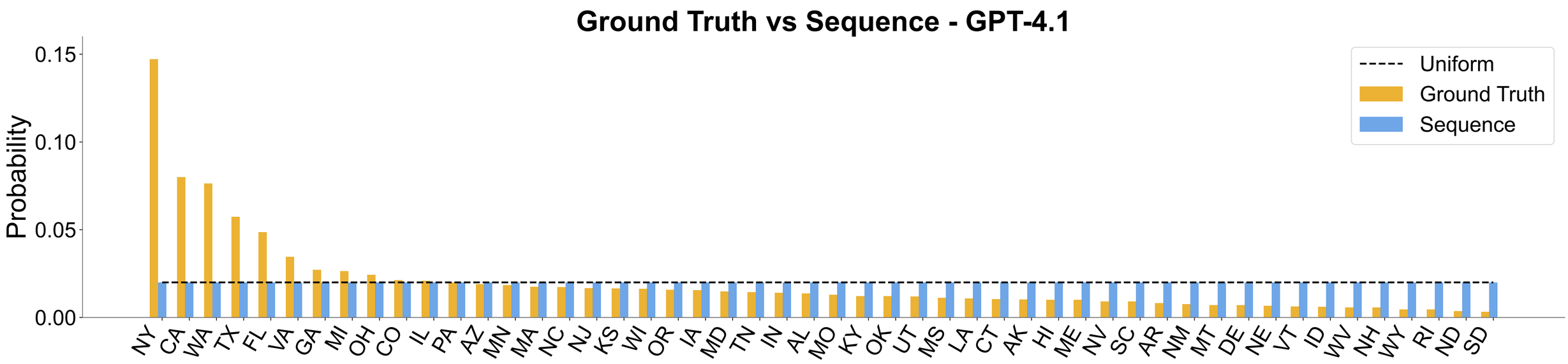
(a)


Ground Truth vs Direct - GPT-4.1
Probability
0.15
0.10
0.05
0.00
Uniform
Ground Truth
Direct
NY CA WA TX FL VA GA MI OH CO IL PA AZ MN MA NC NJ KS WI OR IA MD TN IN AL MO KY OK UT MS LA CT AK HI ME NV SC AR NM MT DE NE VT ID WV NH WY RI ND SD

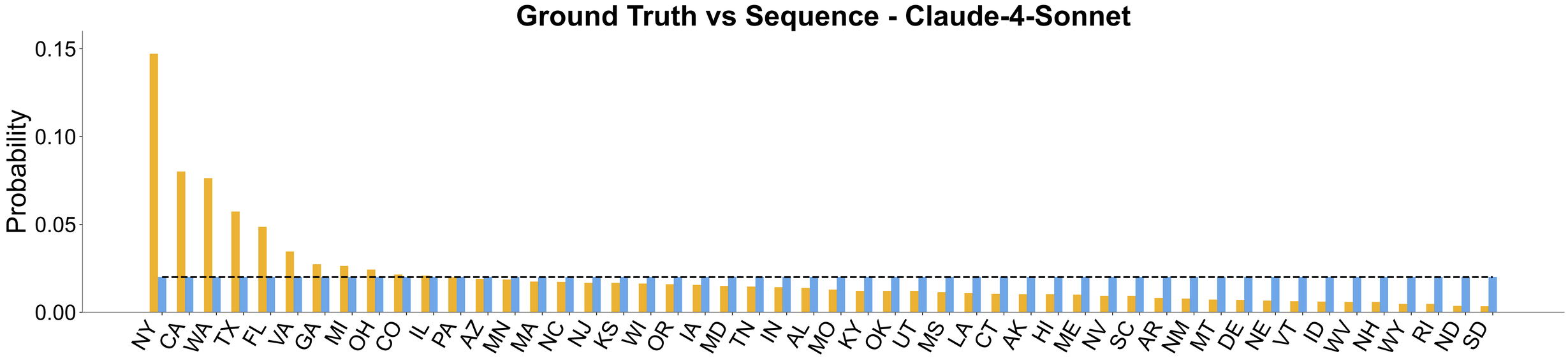
(b)


Ground Truth vs Direct - Claude-4-Sonnet
Probability
0.15
0.10
0.05
0.00
NY CA WA TX FL VA GA MI OH CO IL PA AZ MN MA NC NJ KS WI OR IA MD TN IN AL MO KY OK UT MS LA CT AK HI ME NV SC AR NM MT DE NE VT ID WV NH WY RI ND SD

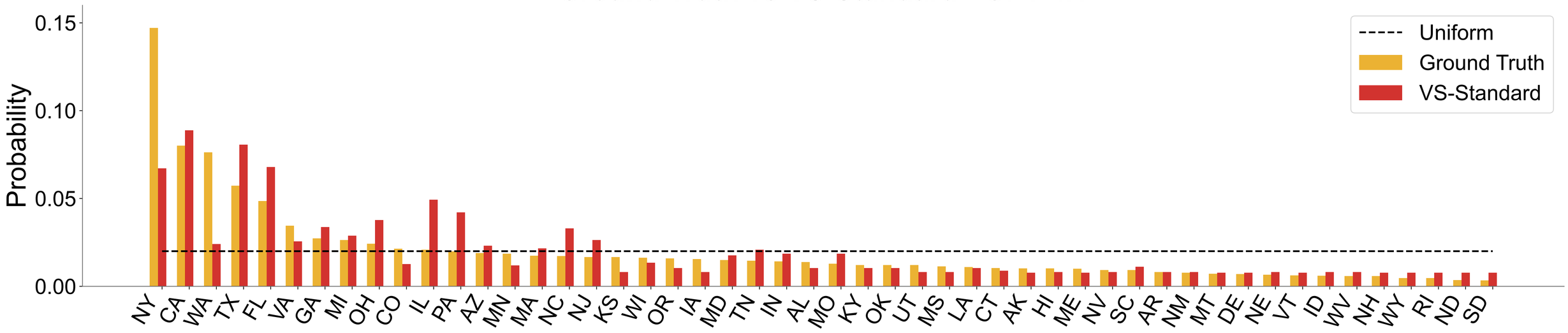
(c)


Ground Truth vs Sequence - GPT-4.1
Probability
0.15
0.10
0.05
0.00
Uniform
Ground Truth
Sequence
NY CA WA TX FL VA GA MI OH CO IL PA AZ MN MA NC NJ KS WI OR IA MD TN IN AL MO KY OK UT MS LA CT AK HI ME NV SC AR NM MT DE NE VT ID WV NH WY RI ND SD

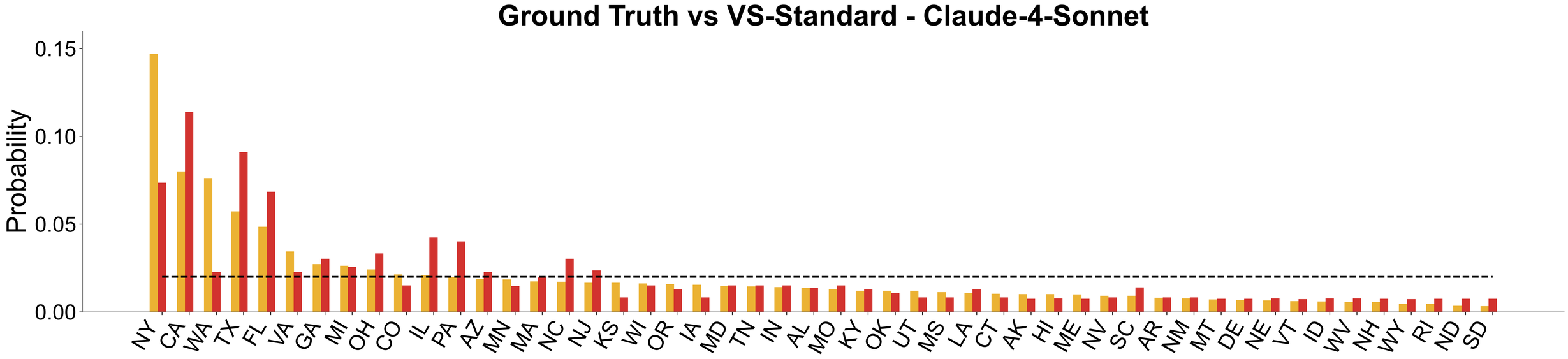
(d)


Ground Truth vs Sequence - Claude-4-Sonnet
Probability
0.15
0.10
0.05
0.00
NY CA WA TX FL VA GA MI OH CO IL PA AZ MN MA NC NJ KS WI OR IA MD TN IN AL MO KY OK UT MS LA CT AK HI ME NV SC AR NM MT DE NE VT ID WV NH WY RI ND SD

(e)


Ground Truth vs VS-Standard - GPT-4.1
Probability
0.15
0.10
0.05
0.00
Uniform
Ground Truth
VS-Standard
NY CA WA TX FL VA GA MI OH CO IL PA AZ MN MA NC NJ KS WI OR IA MD TN IN AL MO KY OK UT MS LA CT AK HI ME NV SC AR NM MT DE NE VT ID WV NH WY RI ND SD

(f)


Ground Truth vs VS-Standard - Claude-4-Sonnet
Probability
0.15
0.10
0.05
0.00
NY CA WA TX FL VA GA MI OH CO IL PA AZ MN MA NC NJ KS WI OR IA MD TN IN AL MO KY OK UT MS LA CT AK HI ME NV SC AR NM MT DE NE VT ID WV NH WY RI ND SD