**Graph 3: Capabilities with Largest Accuracy-Robustness Discrepancy**
**Question: Which capabilities show accuracy-robustness gaps?**

LARGEST GAP: 0.487
Capability: Conducting comprehensive quantitative reasoning

Legend: Traditional Accuracy, DOVE Robustness

- economic data, principles, policies — Gap: 0.247
- data methodologies validity, accuracy, — Gap: 0.252
- moral, ethical, social, legal — Gap: 0.259
- moral implications based ethical, — Gap: 0.264
- chemical physical properties interactions — Gap: 0.282
- Analyzing and synthesizing electromagnetic — Gap: 0.286
- chemical reactions, interactions, properties — Gap: 0.310
- complex scientific principles systems — Gap: 0.320
- scientific wave-related phenomena principles — Gap: 0.333
- optimizing complex scientific thermodynamic — Gap: 0.345
- Enhancing statistical probabilistic reasoning — Gap: 0.362
- Utilizing mathematical reasoning problem-solving — Gap: 0.375
- applying advanced scientific mathematical — Gap: 0.401
- Developing quantitative reasoning analytical — Gap: 0.411
- Conducting comprehensive quantitative reasoning — Gap: 0.487

Score