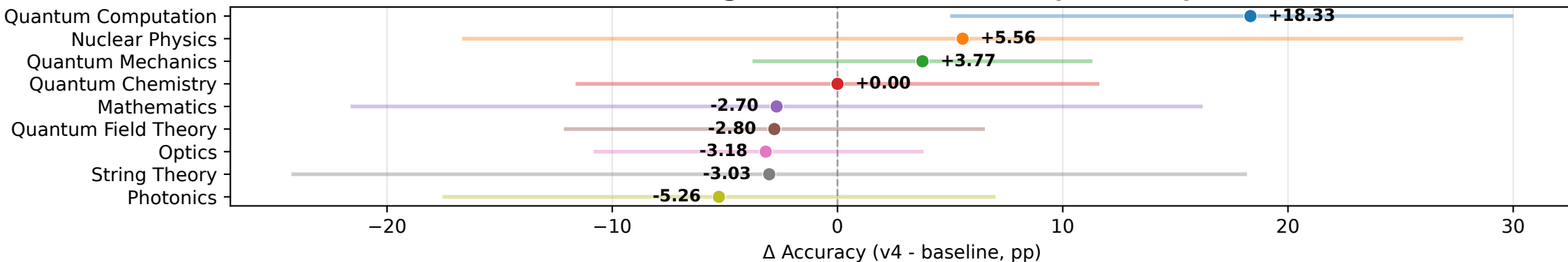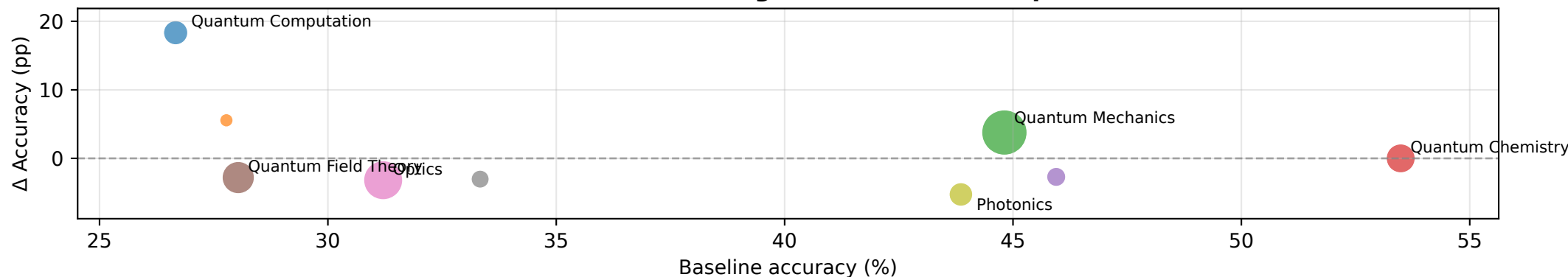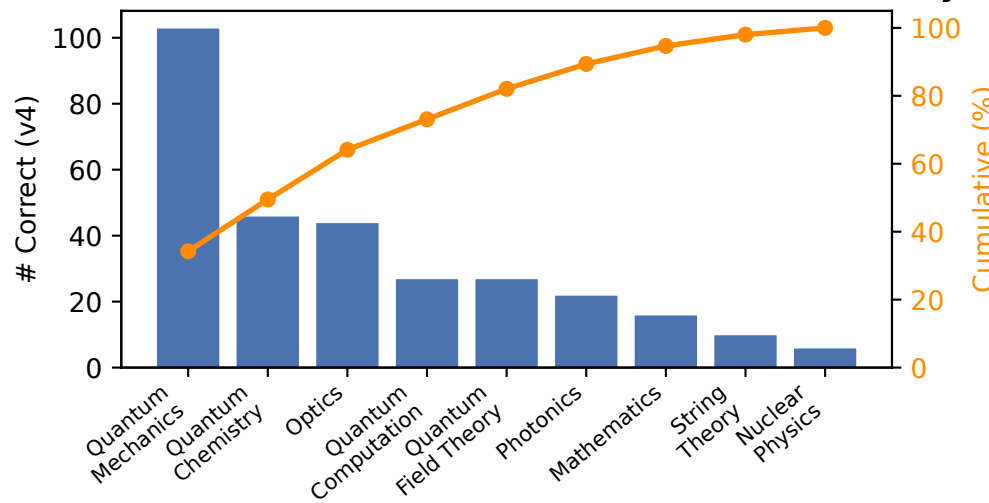**(a) Subdomain gains/losses with bootstrap 95% CI (paired)**

(b) Gain vs baseline strength (bubble size ∝ #questions)

(c) Pareto: subdomain contribution to v4 accuracy

(d) Δpp by question type × subdomain (n≥2)