

Experiments on the Test Case Length

in Specification Based Test Case Generation

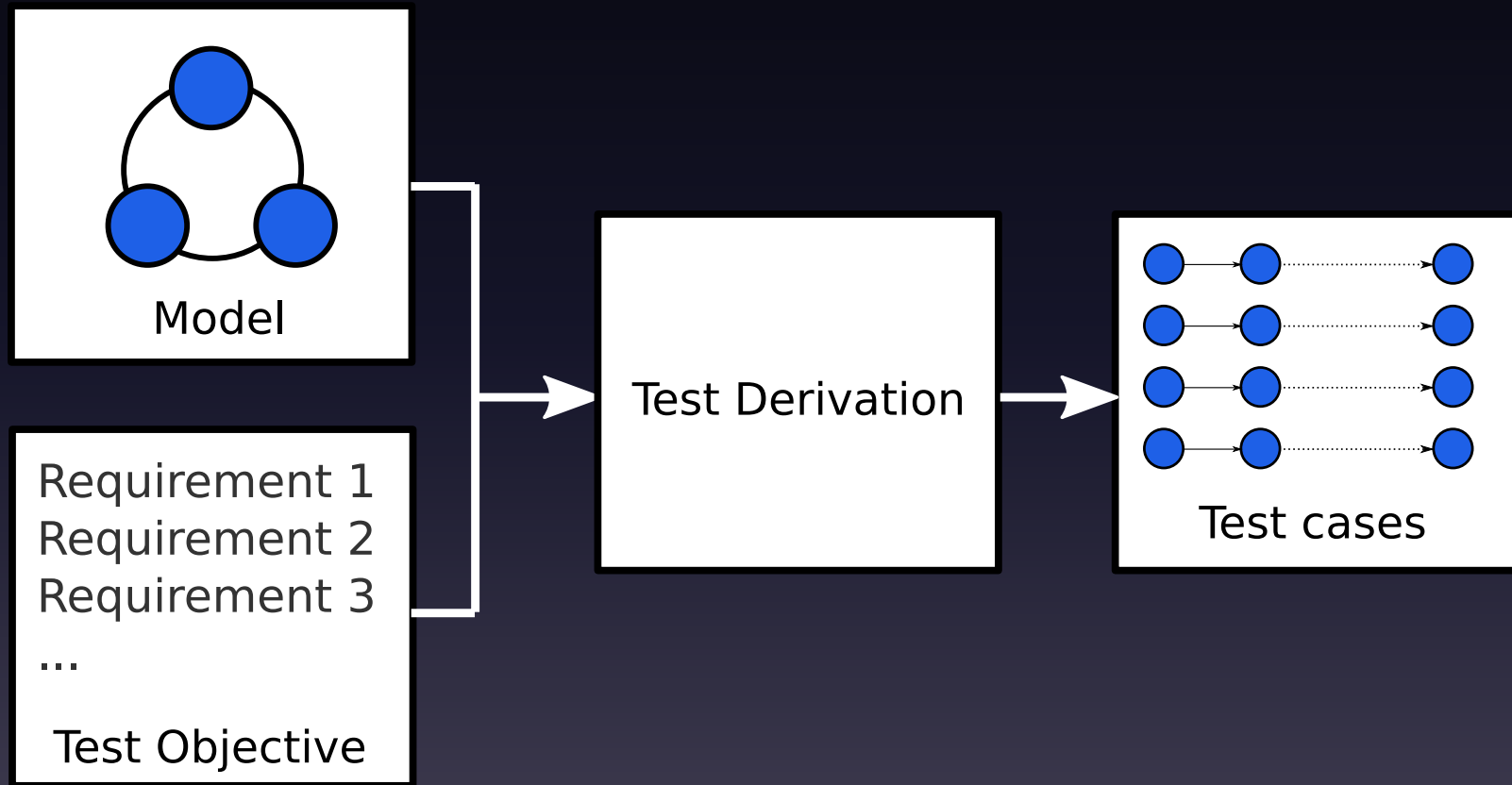
Gordon Fraser Angelo Gargantini

Graz University of Technology Austria

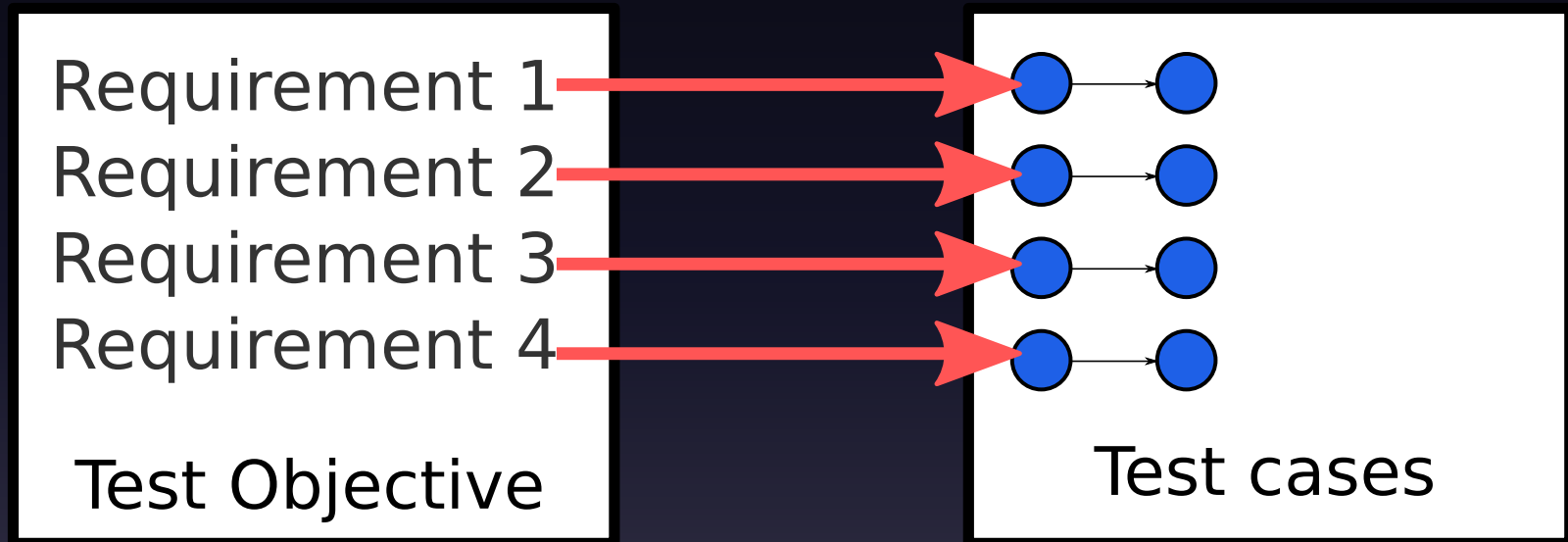
May 18, 2009

Introduction

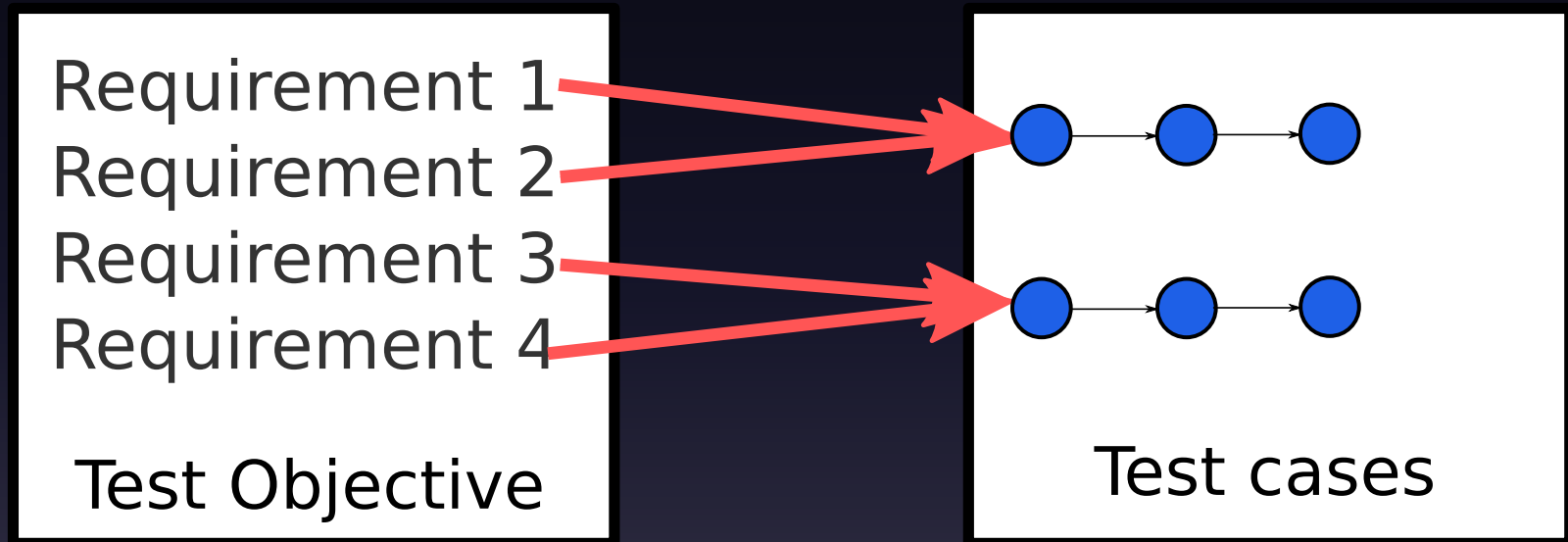
Specification based testing



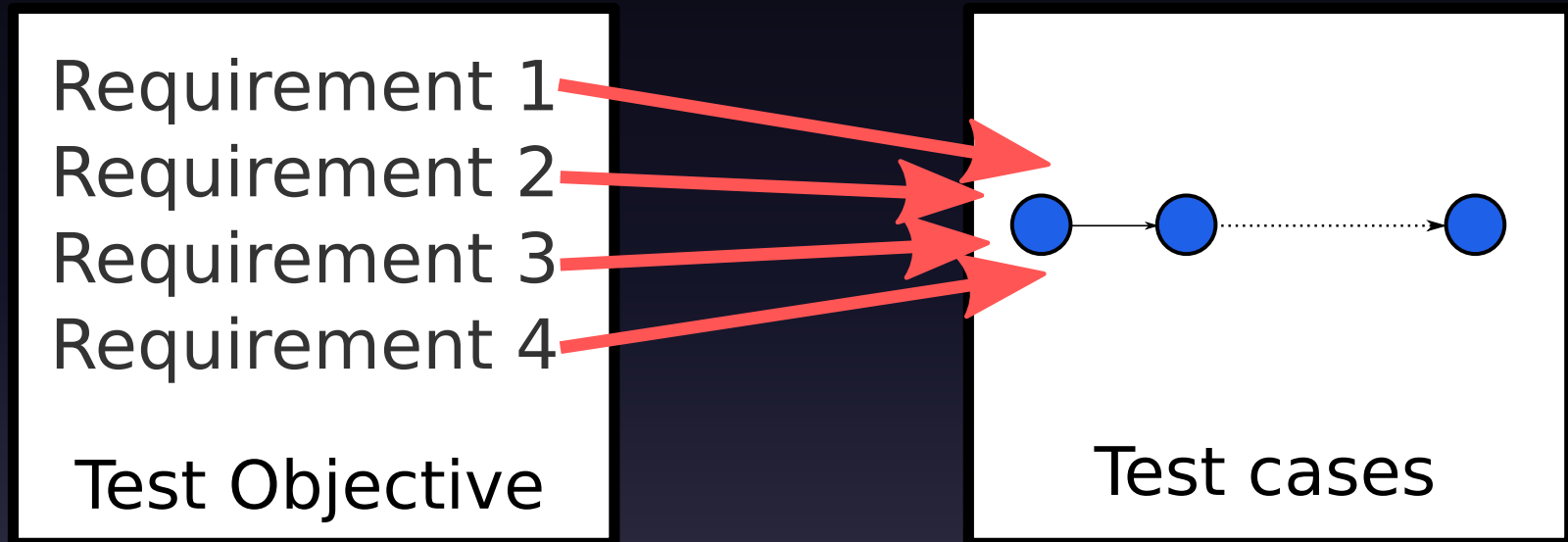
Motivation



Motivation



Motivation



Motivation

What are the effects of the test case length?

- Costs
 - Test case generation costs
 - Test execution costs
- Fault detection ability
 - Coverage
 - Mutation score

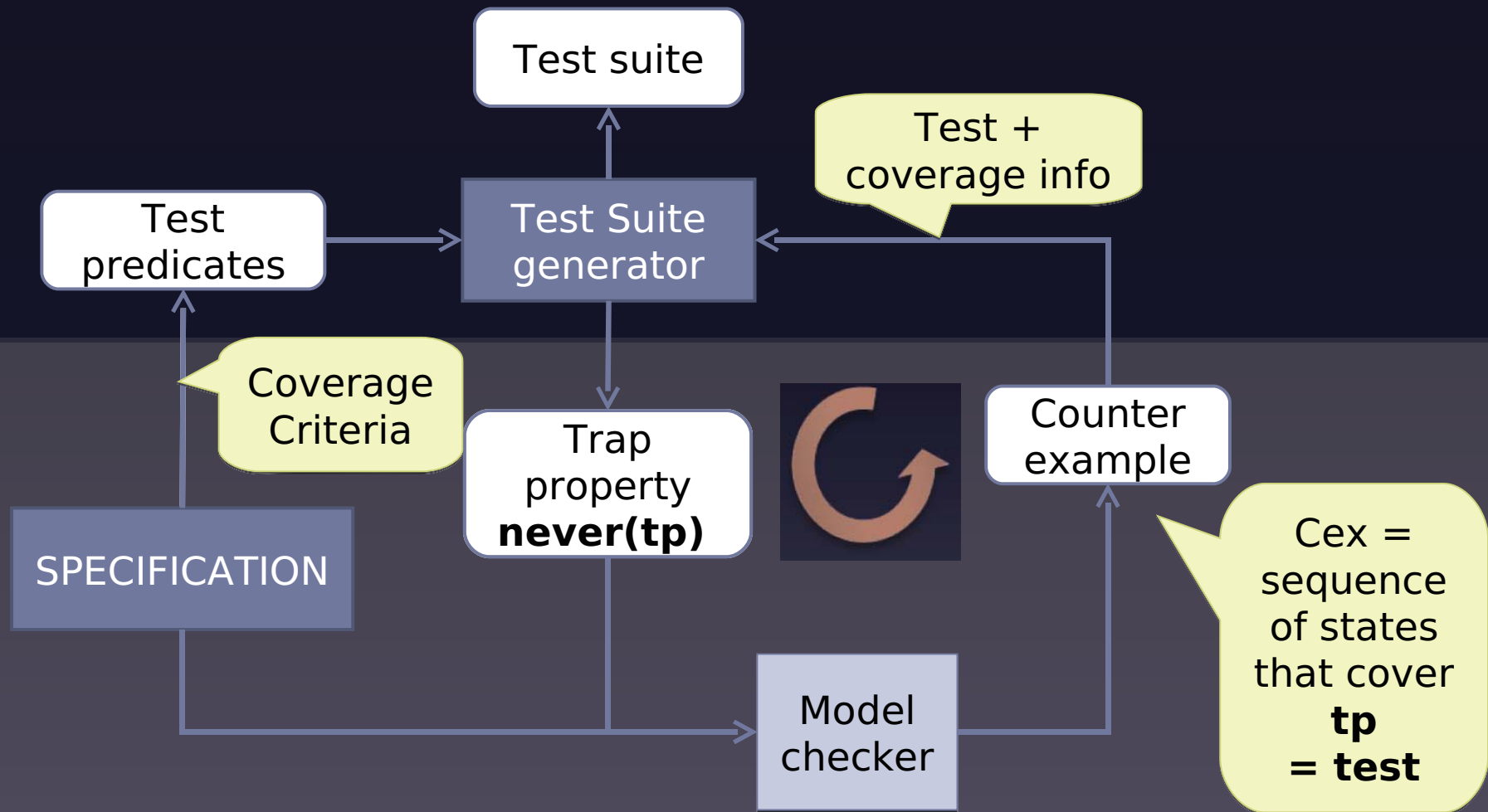
Motivation

What are the effects of the test case length?

- Costs
 - Test case generation costs
 - Test execution costs
- Fault detection ability
 - Coverage
 - Mutation score
- What is preferable - many short test cases or fewer longer test cases?

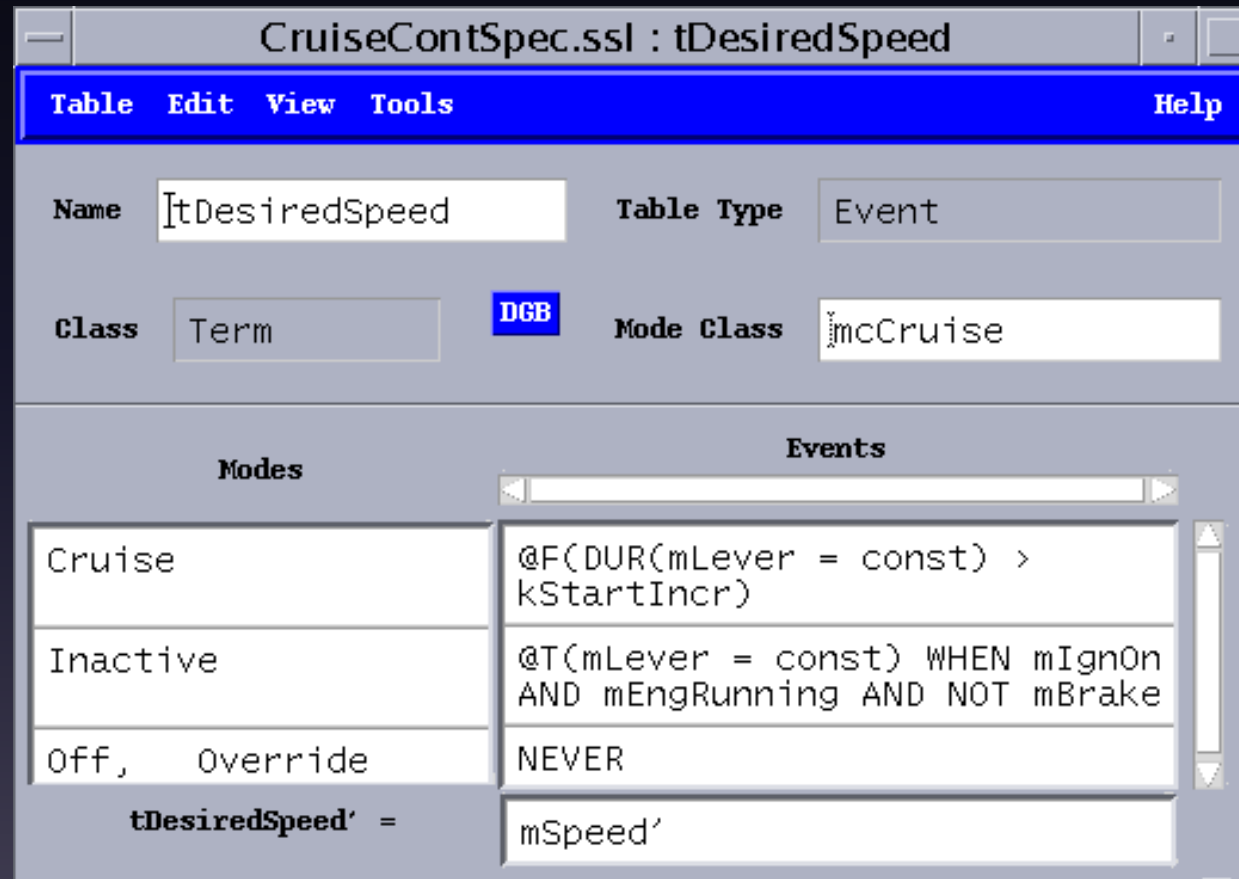
Experimental Setup

Test case generation with model checkers



Experimental Setup

- Specifications:
 - Safety Injection System (SIS)
 - Cruise Control



Experimental Setup

- Specifications:
 - Safety Injection System (SIS)
 - Cruise Control
- Coverage criteria:
 - Mutation, MCDC, Transition Pair, Pairwise
 - MCDC
 - $\mathbf{AG} \neg(\alpha \wedge (\gamma_{C, False} \oplus \gamma_{C, True}) \wedge C)$
 - $\mathbf{AG} \neg(\alpha \wedge (\gamma_{C, False} \oplus \gamma_{C, True}) \wedge \neg C)$
 - Transition Pair
 - $\mathbf{AG} \neg((\alpha_1 \wedge \gamma_1) \wedge \mathbf{EX} (\alpha_2 \wedge \gamma_2))$
 - Mutation
 - $\mathbf{AG} \neg(\alpha \wedge (M \oplus \gamma))$
 - Pairwise
 - $\mathbf{AG} \neg(V_1 = v_1 \wedge V_2 = v_2)$

Experimental Setup

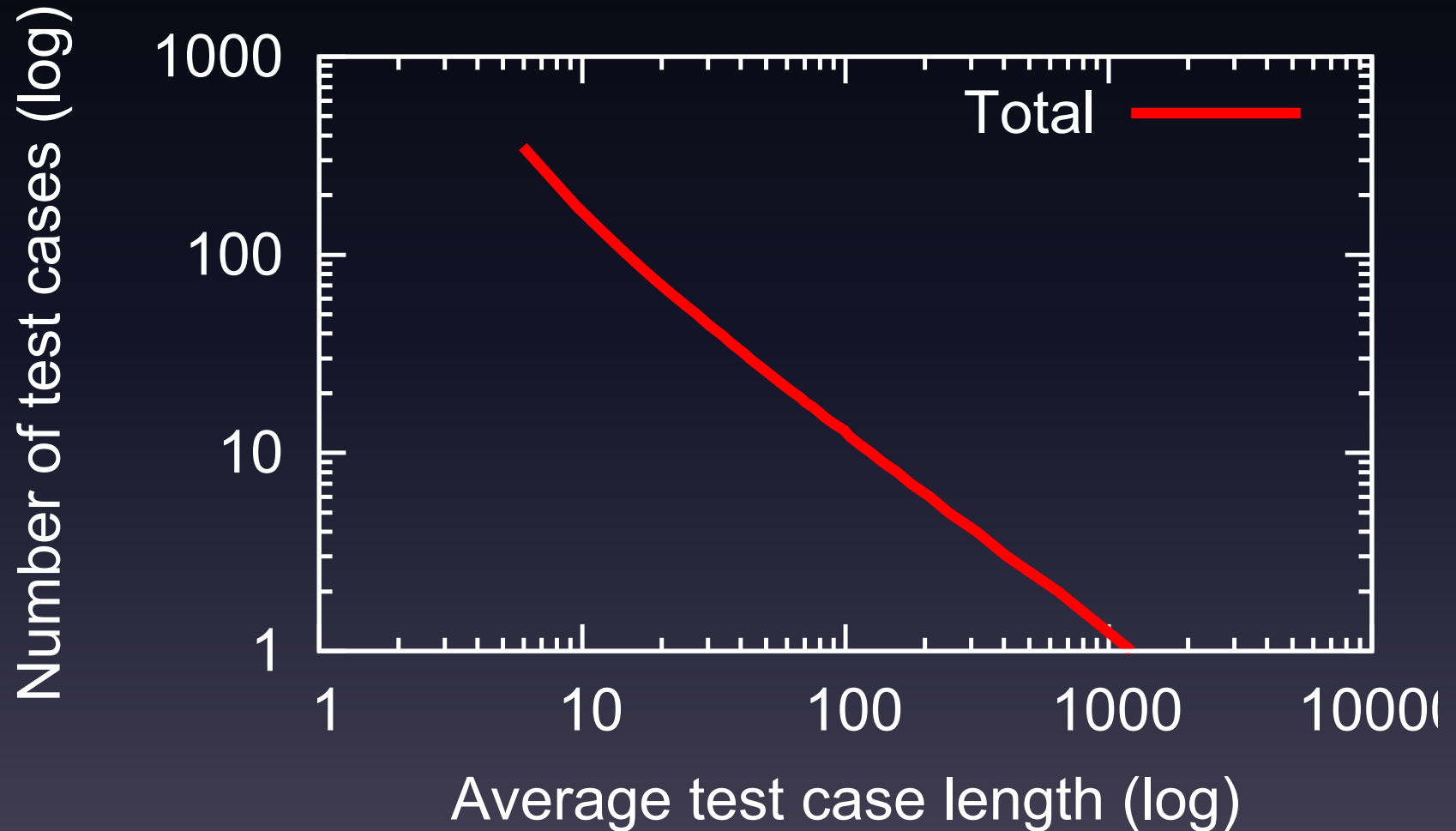
- Specifications:
 - Safety Injection System (SIS)
 - Cruise Control
- Coverage criteria:
 - Mutation, MCDC, Transition Pair, Pairwise
- Model Checker:
 - NuSMV, symbolic model checker

Experimental Setup

- Specifications:
 - Safety Injection System (SIS)
 - Cruise Control
- Coverage criteria:
 - Mutation, MCDC, Transition Pair, Pairwise
- Model Checker:
 - NuSMV, symbolic model checker
- For each spec, for each criterion:
 - Create 1 test case for 1 trap property
 - Create 1 test case for 2 trap properties
 - ...
 - Create 1 test case for all trap properties

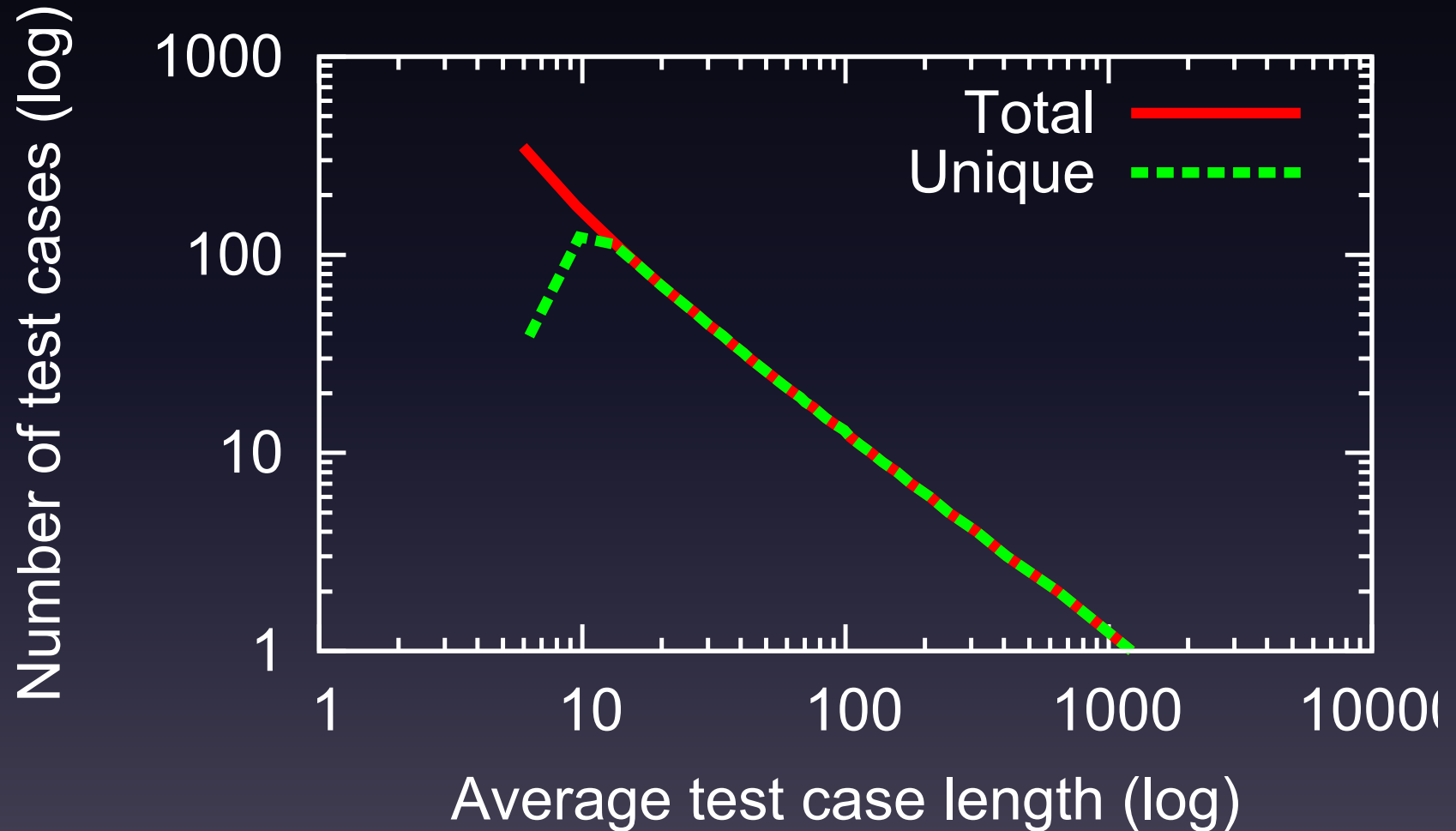
Effects on the Test Suite Size

Cruise, Mutation



Effects on the Test Suite Size

Cruise, Mutation



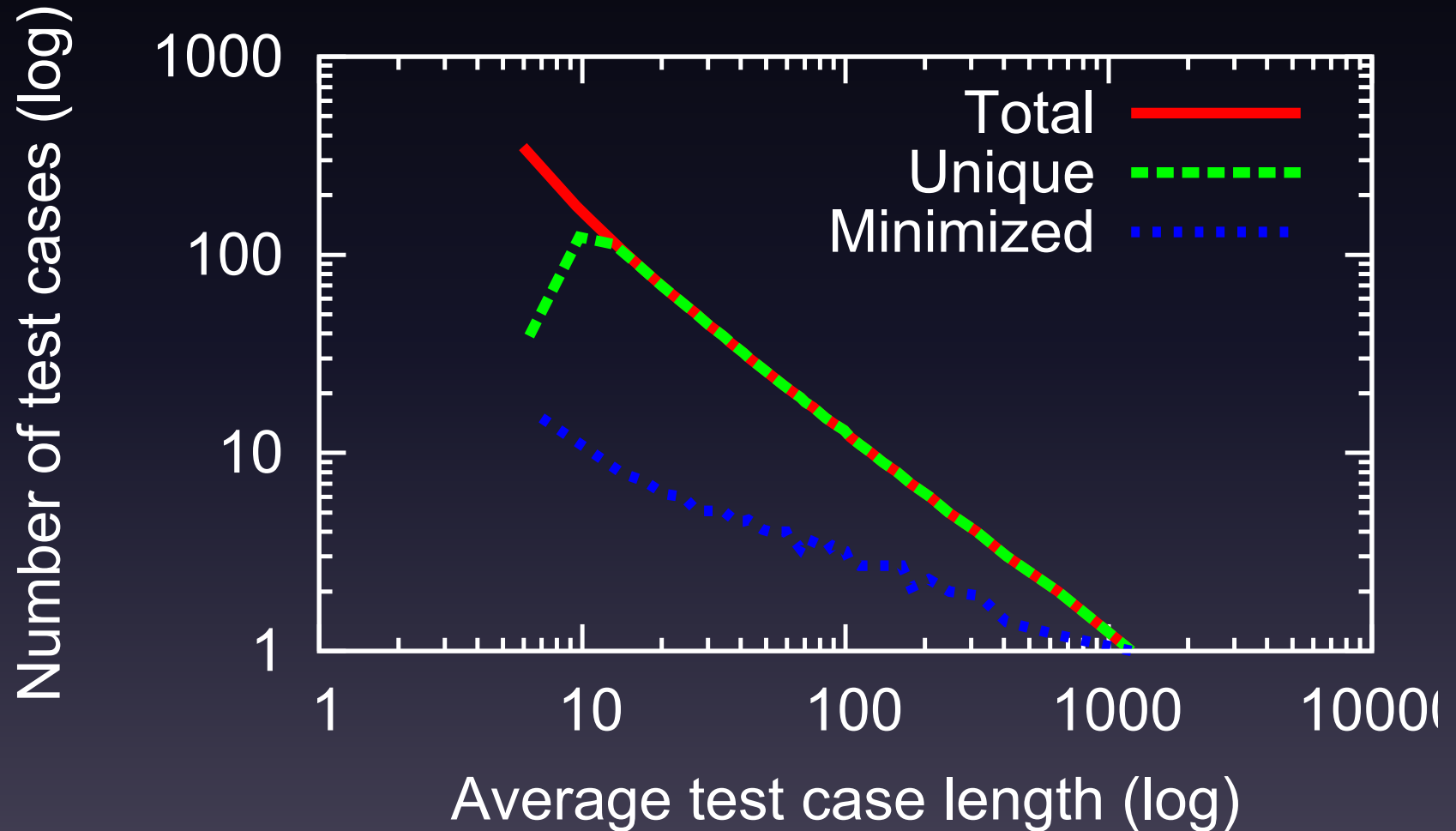
Effects on the Test Suite Size

Minimization

- Minimal = no test case can be dropped wrt. test objective
- Many different heuristics have been proposed
- Applied heuristic:
 - Order test cases by number of satisfied test requirements
 - Pick test case that satisfies most test requirements
 - Remove covered test requirements

Effects on the Test Suite Size

Cruise, Mutation



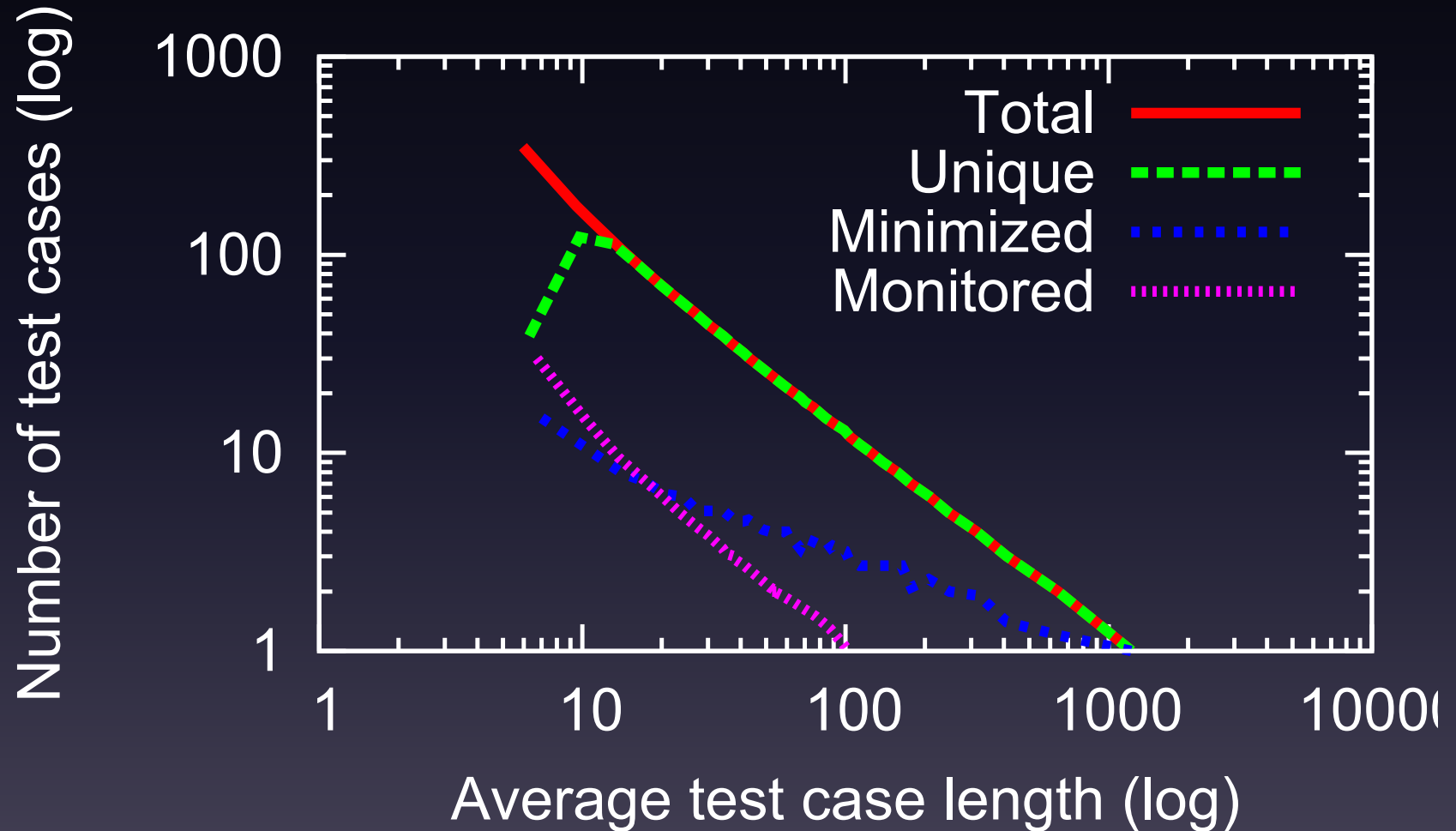
Effects on the Test Suite Size

Monitoring

- Switch between test generation and test requirement checking
- Drop already covered test requirements
- Generate new test cases only for uncovered test requirements
- Does not guarantee minimality

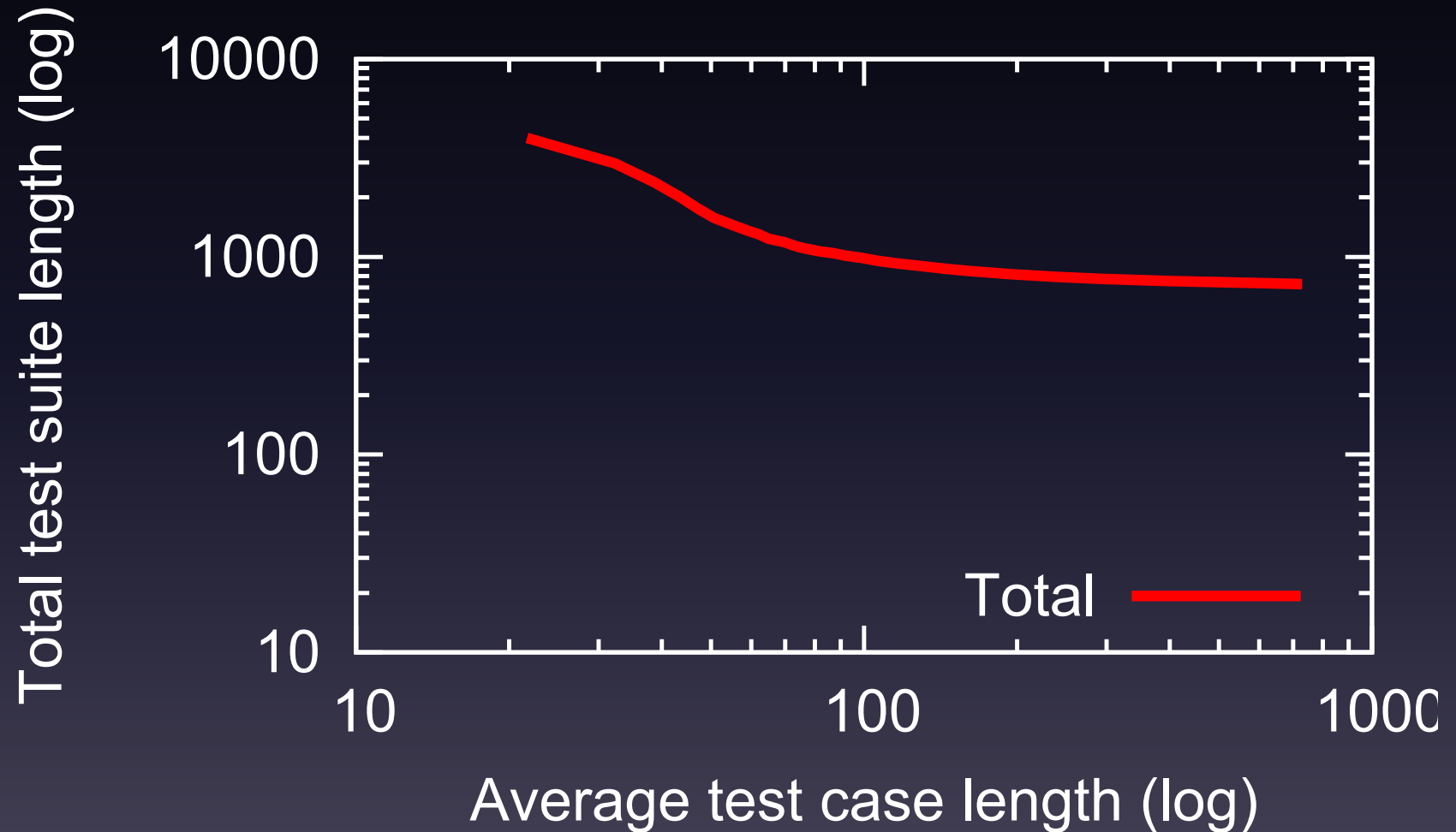
Effects on the Test Suite Size

Cruise, Mutation



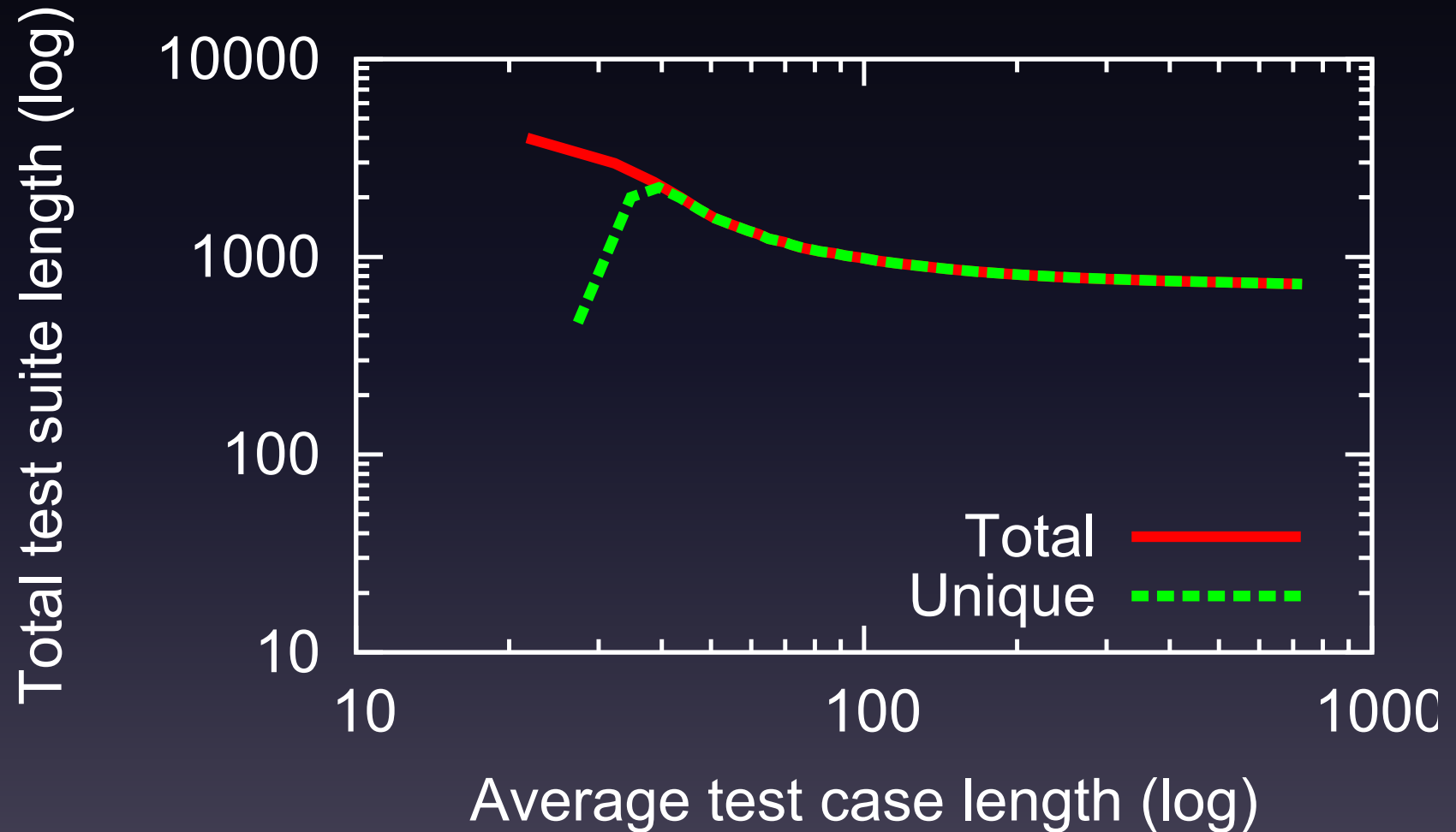
Effects on the Test Suite Length

SIS, Mutation



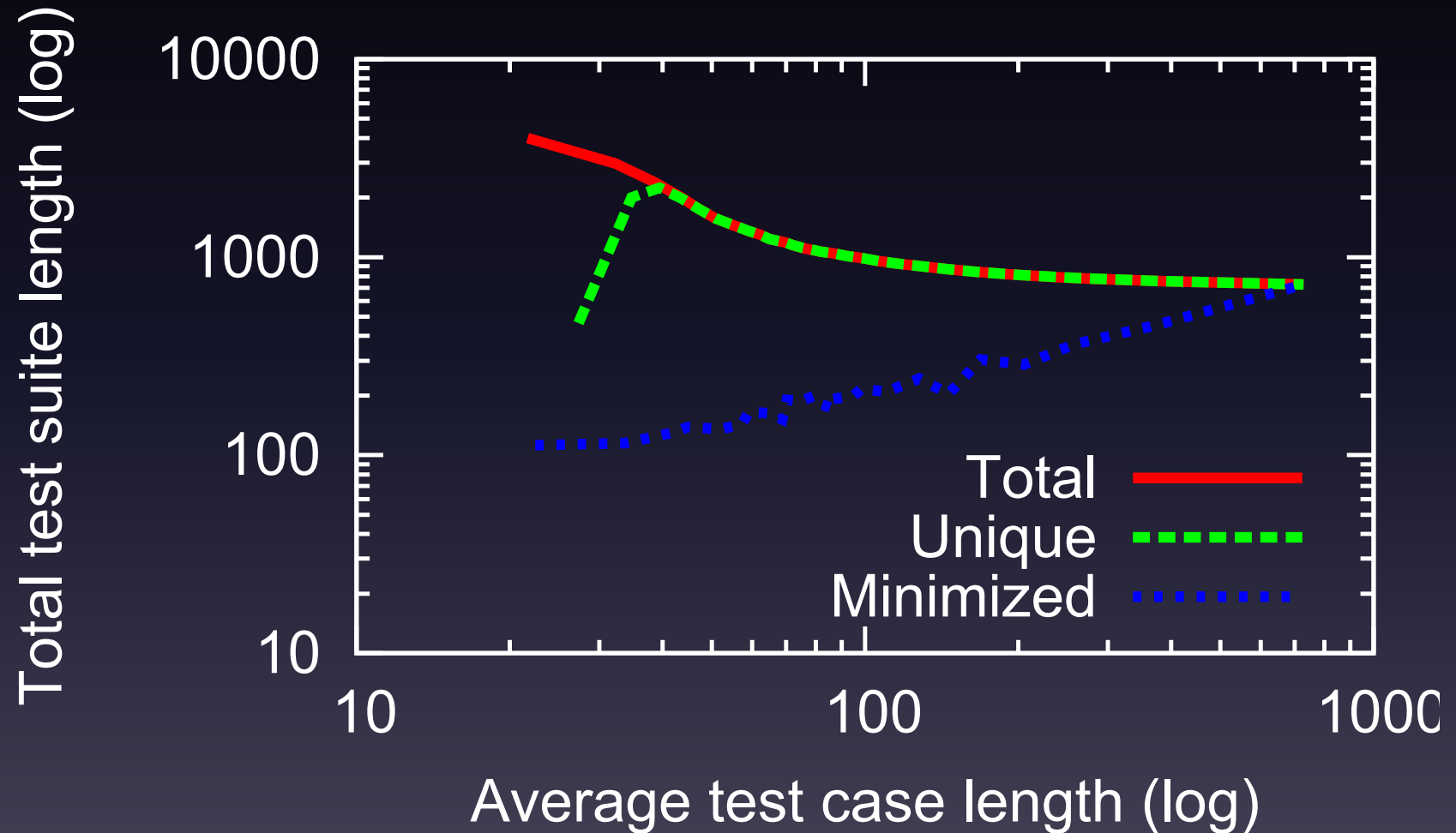
Effects on the Test Suite Length

SIS, Mutation



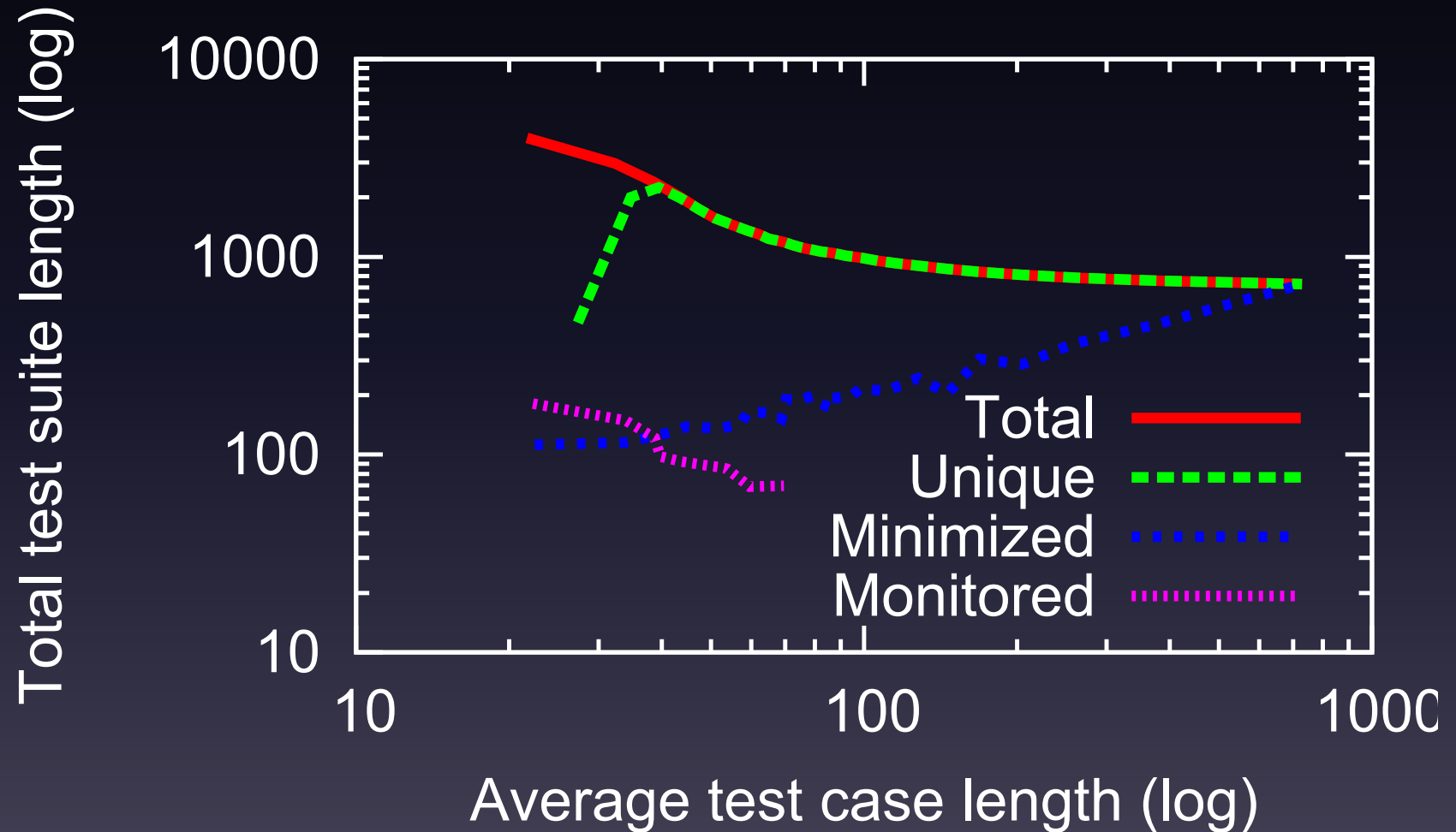
Effects on the Test Suite Length

SIS, Mutation



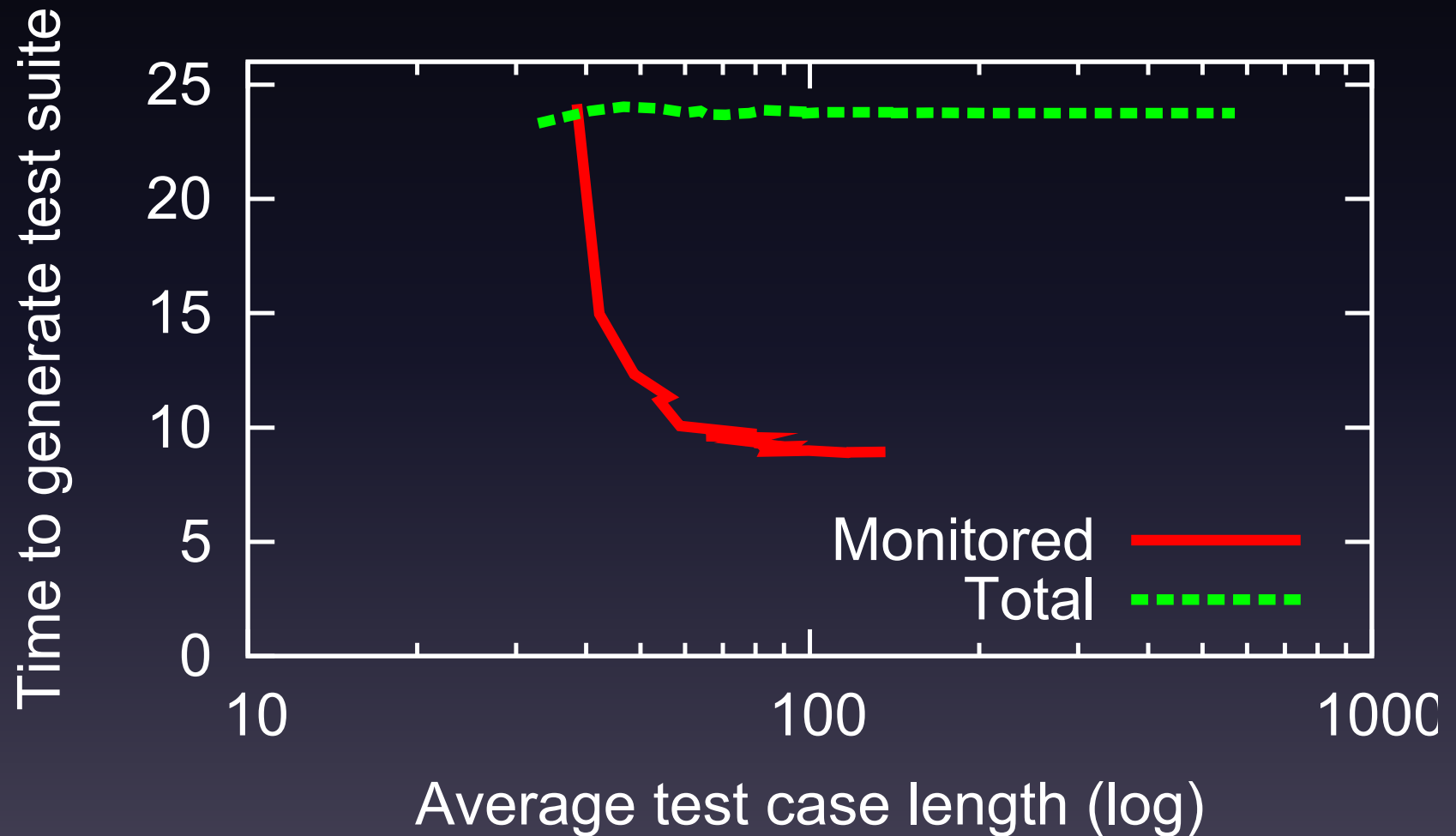
Effects on the Test Suite Length

SIS, Mutation



Effects on the Test Generation Time

SIS, Transition Pair



Effects on the Costs

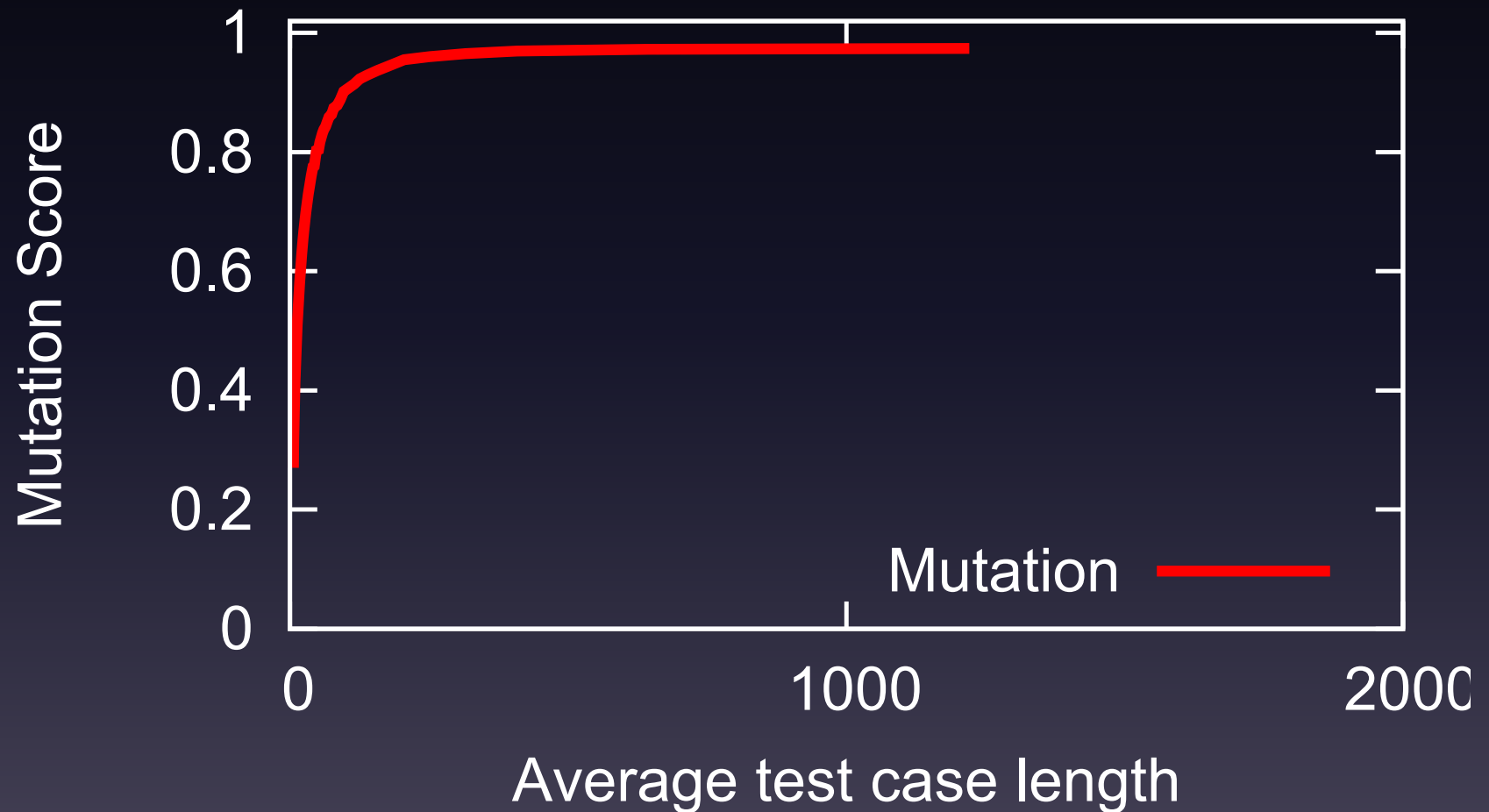
- Test generation
 - Reduction for monitoring + long tests
 - Long tests avoid duplicates → less post-processing
 - Other than that, little effect

Effects on the Costs

- Test generation
 - Reduction for monitoring + long tests
 - Long tests avoid duplicates → less post-processing
 - Other than that, little effect
- Test execution
 - Smallest total length: Long + monitoring
 - Increase in number of unique tests
 - Negative effect on minimization

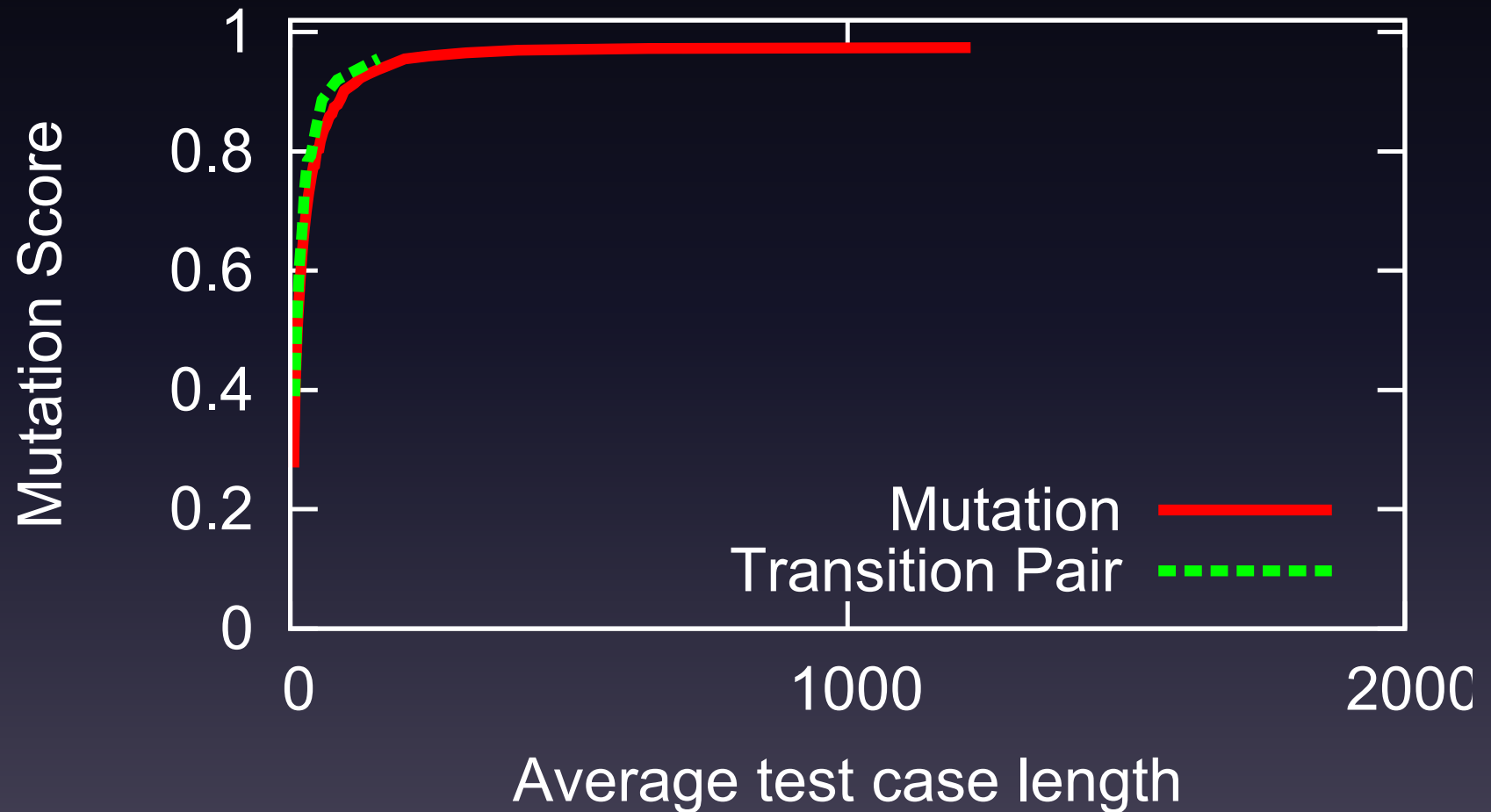
Effects on the Mutation Score

Cruise, Individual test cases



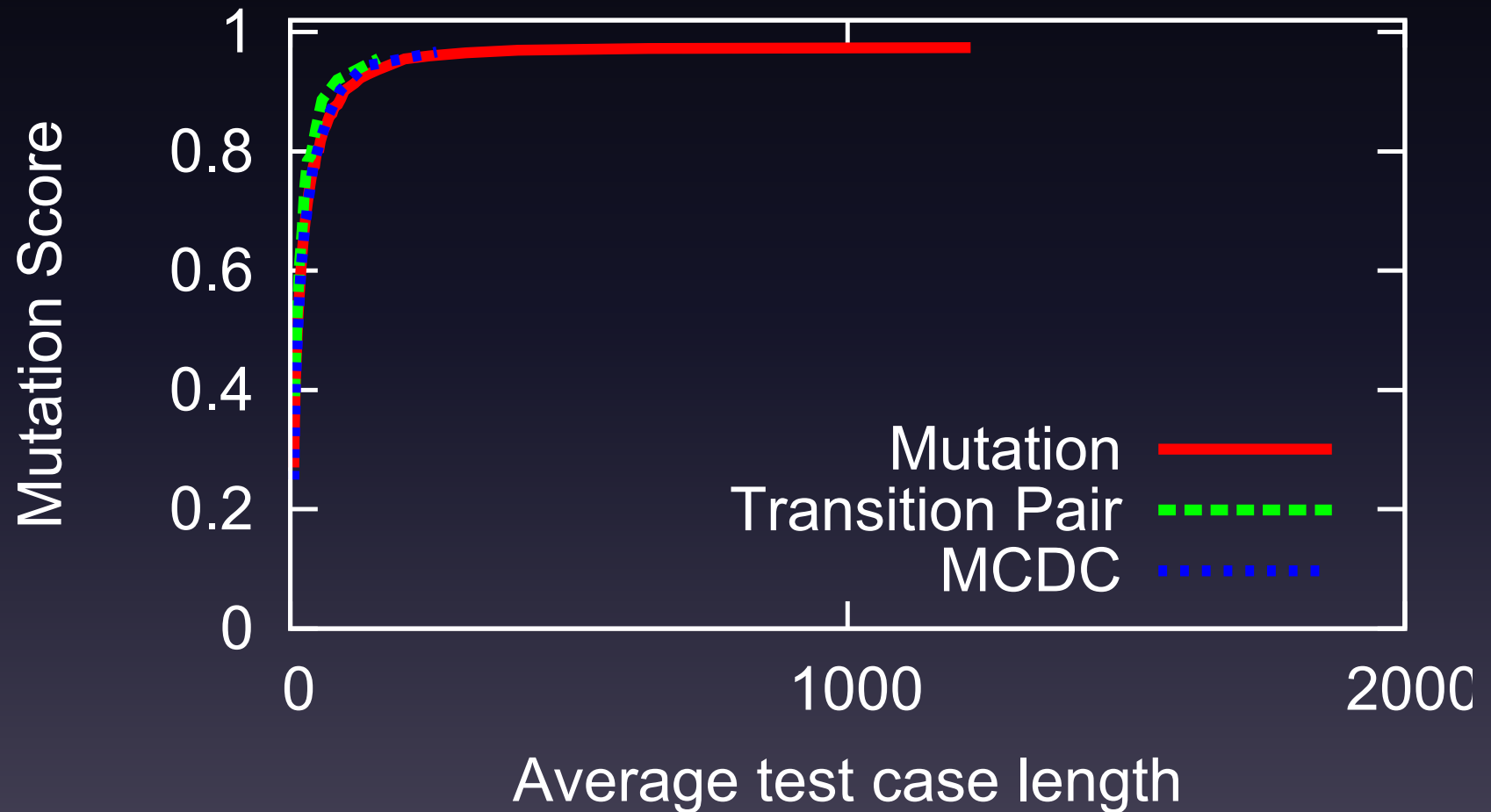
Effects on the Mutation Score

Cruise, Individual test cases



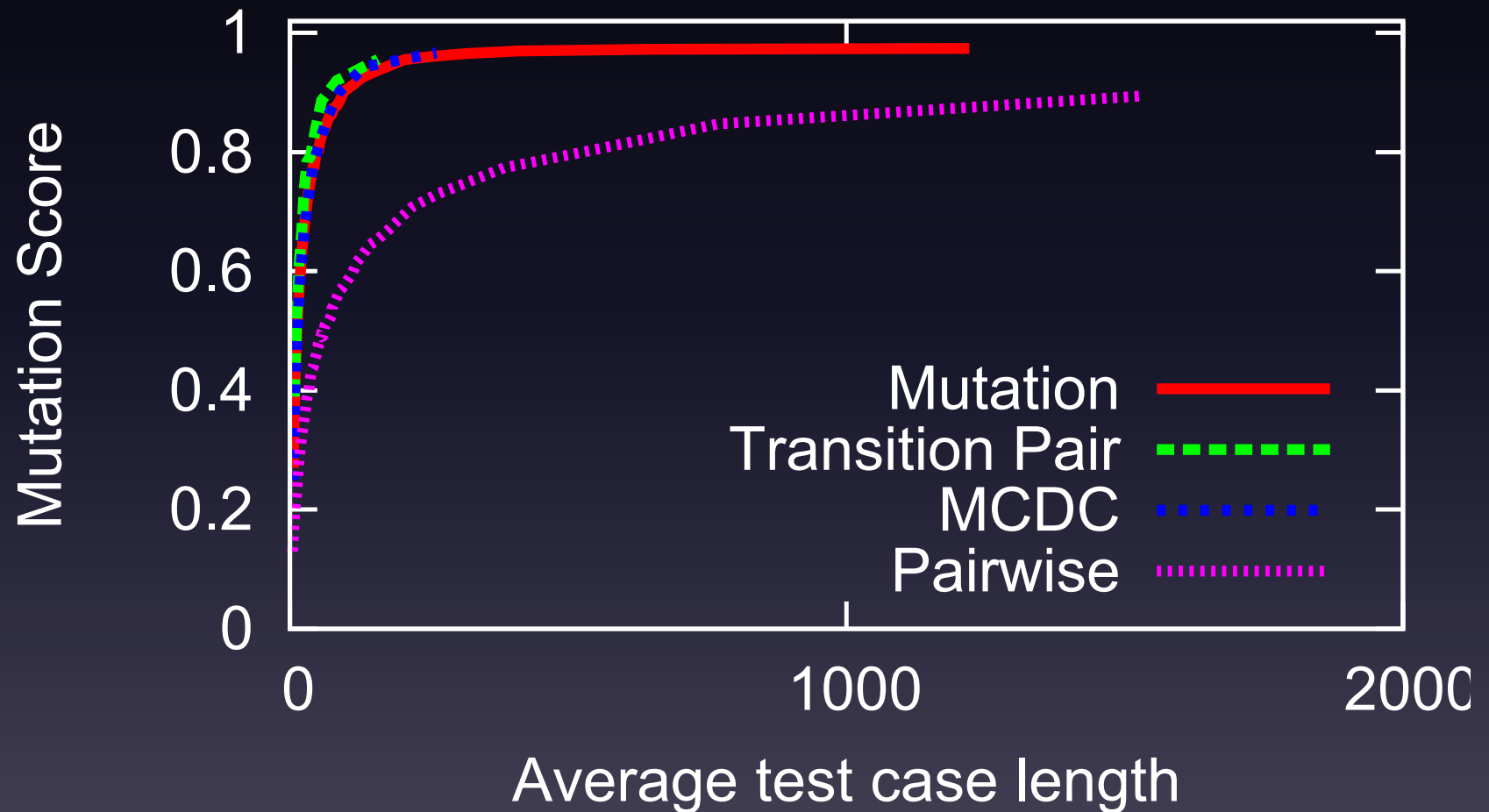
Effects on the Mutation Score

Cruise, Individual test cases



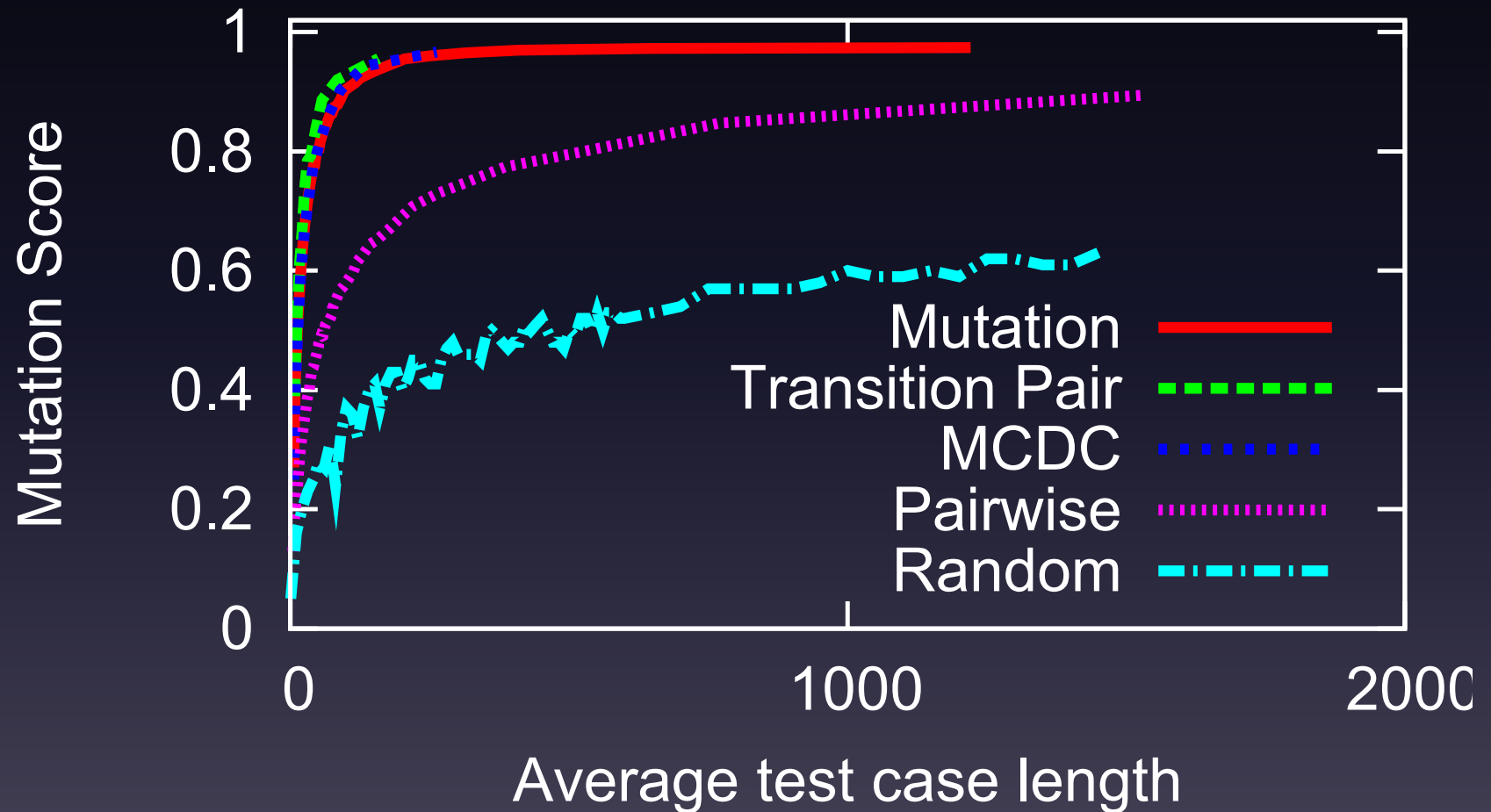
Effects on the Mutation Score

Cruise, Individual test cases



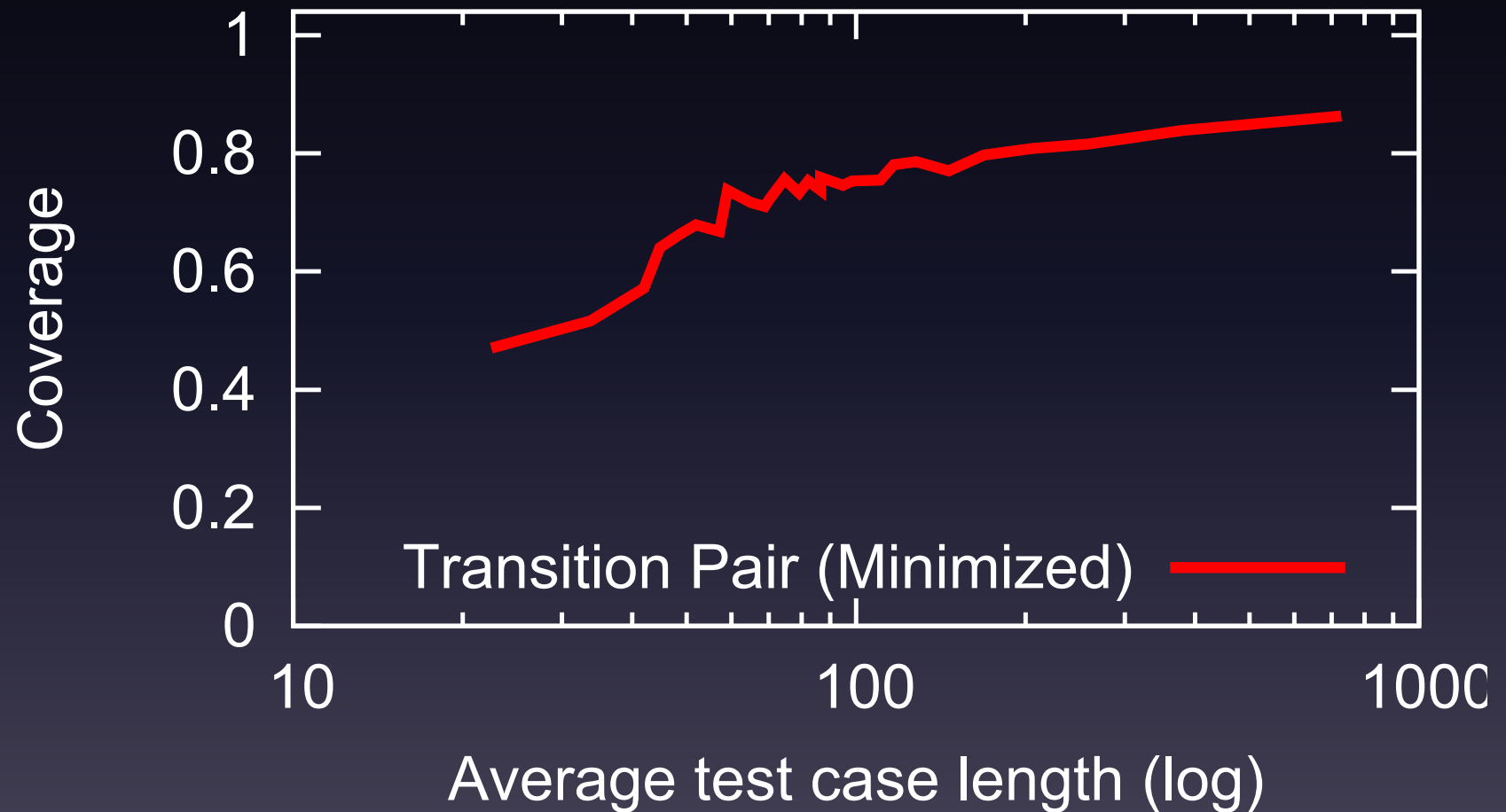
Effects on the Mutation Score

Cruise, Individual test cases



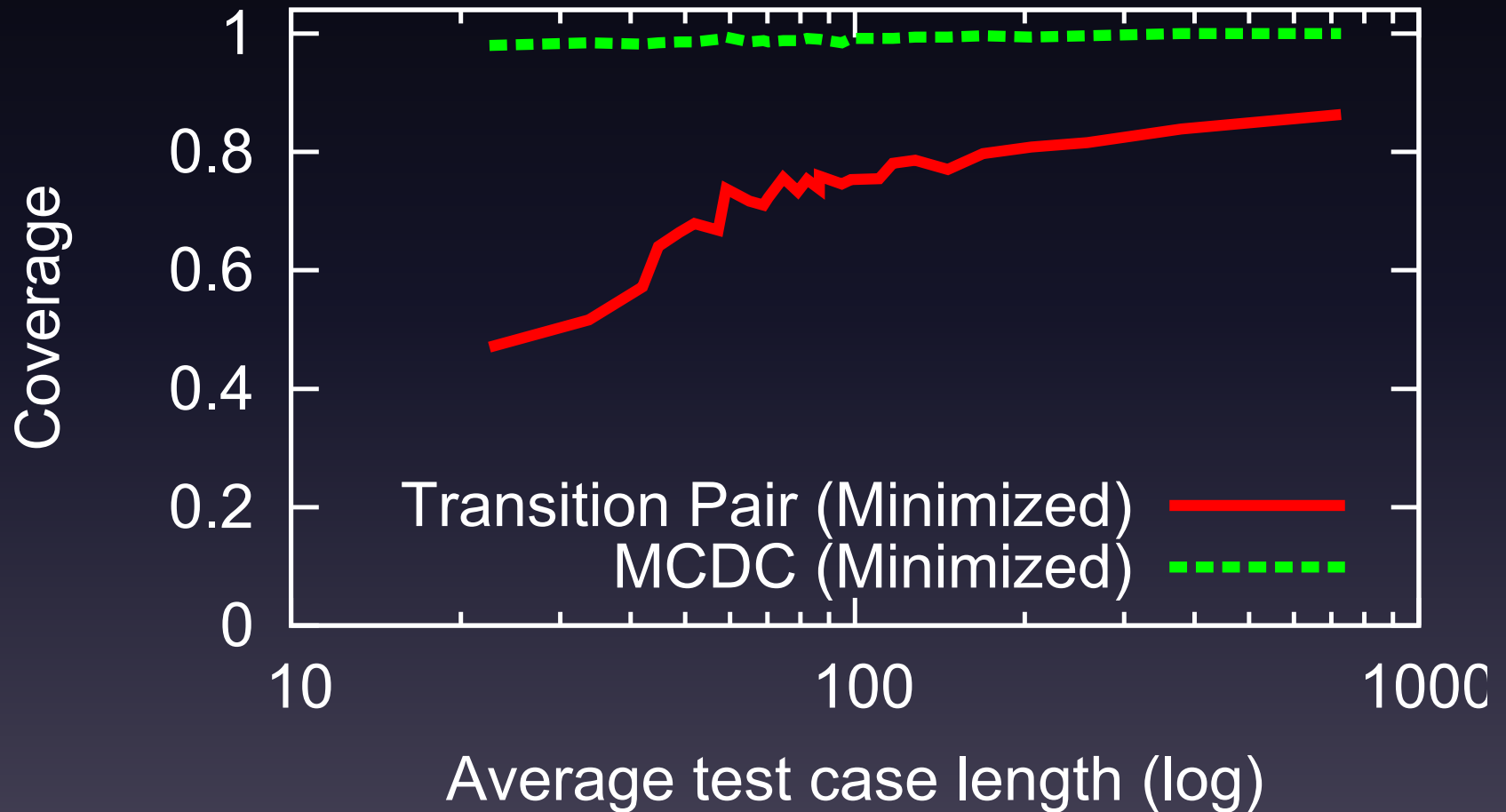
Effects on the Mutation Score

SIS, Test suites



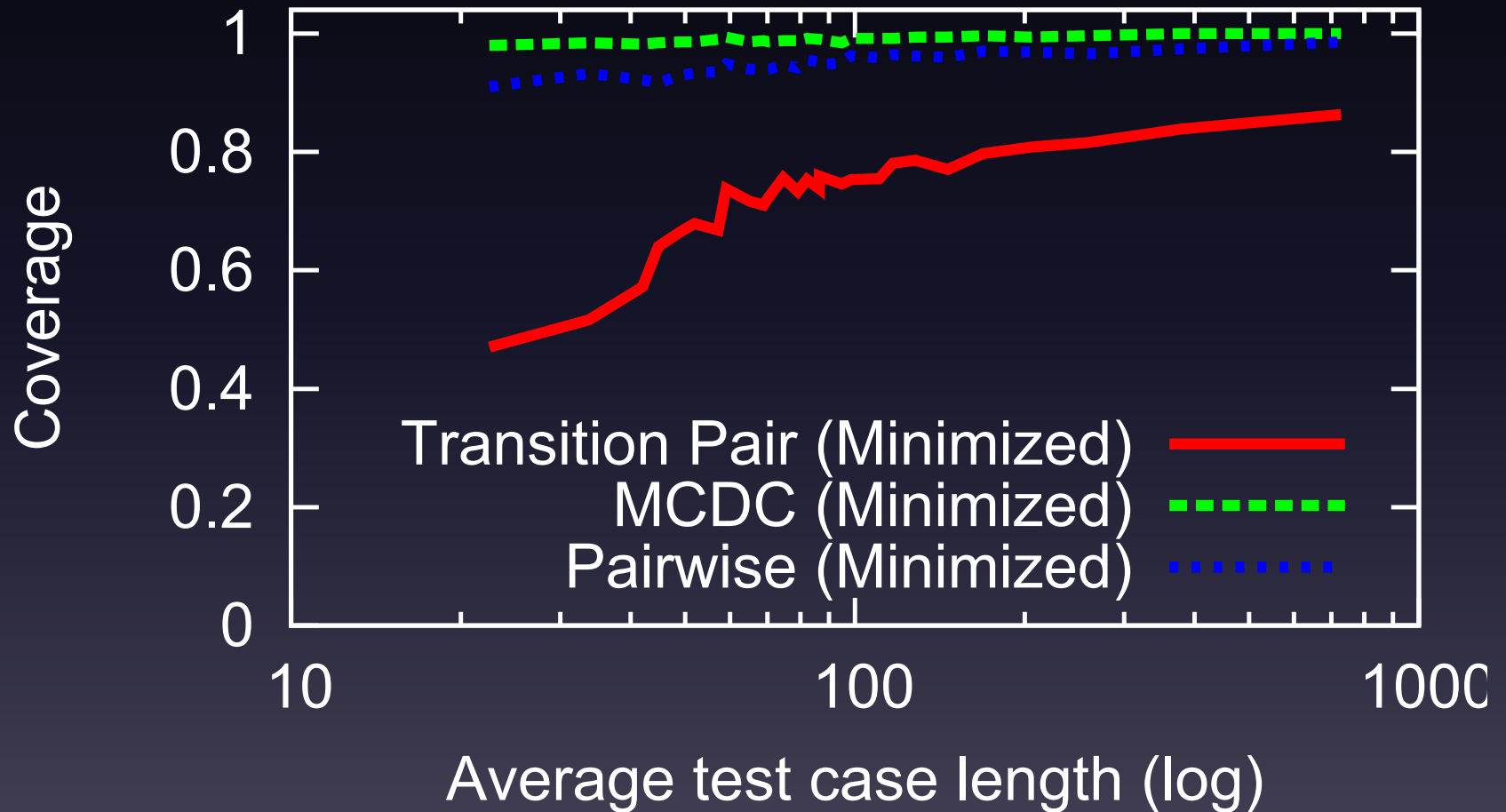
Effects on the Mutation Score

SIS, Test suites



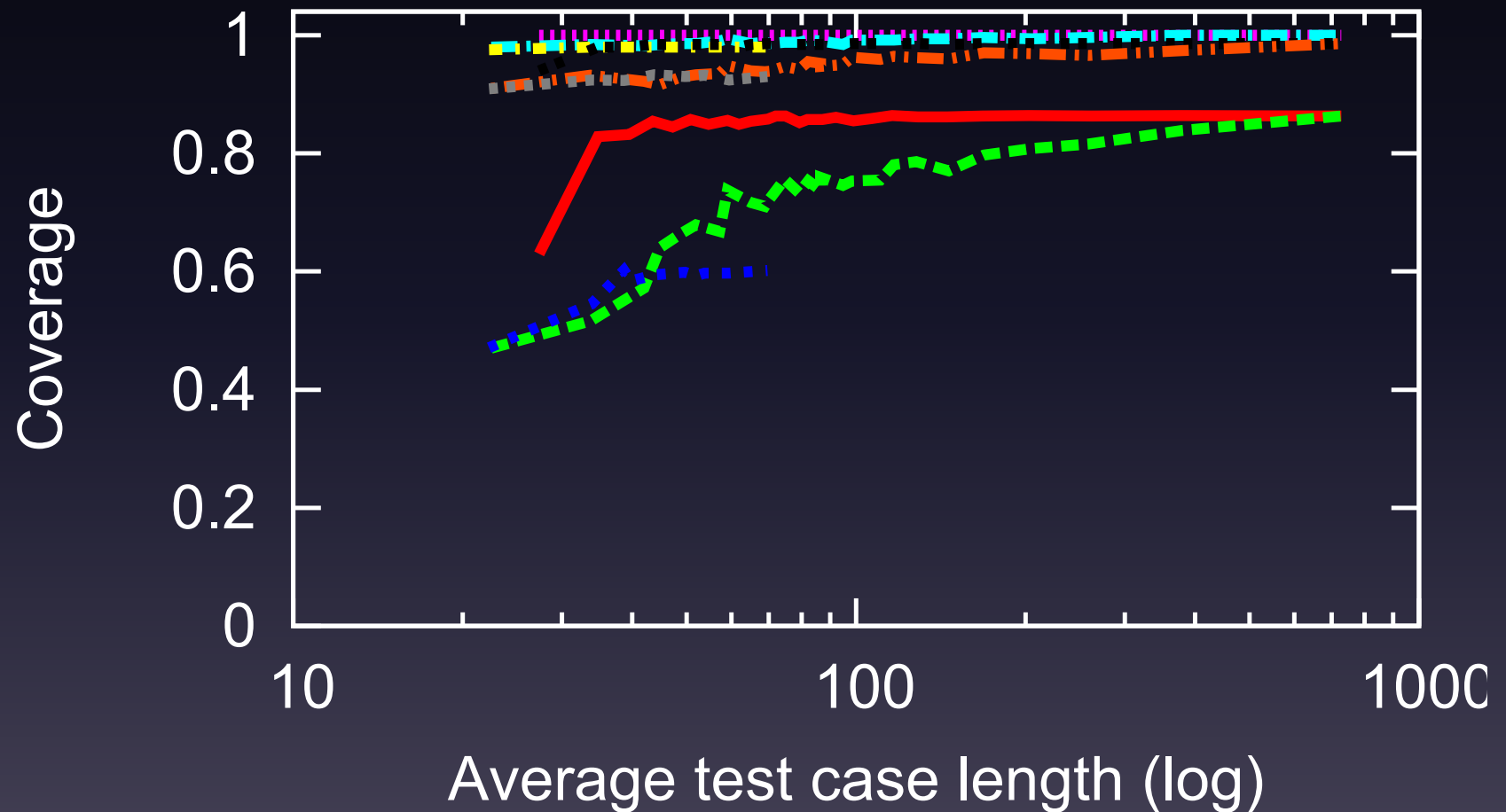
Effects on the Mutation Score

SIS, Test suites



Effects on the Mutation Score

SIS, Test suites



Effects on the Mutation Score

- Longer test cases → slightly higher mutation score+coverage
- Very long test cases → little improvement over not so long test cases
- No decrease in mutation score even when size/length are reduced (e.g., monitoring + long)

Threats to Validity

- Restrictions
 - Determinism
 - Observability
 - Infinite execution
 - Fully connected
- Influence of test generation method?
- Effects of larger specifications?
- Effects on real faults?

Conclusions

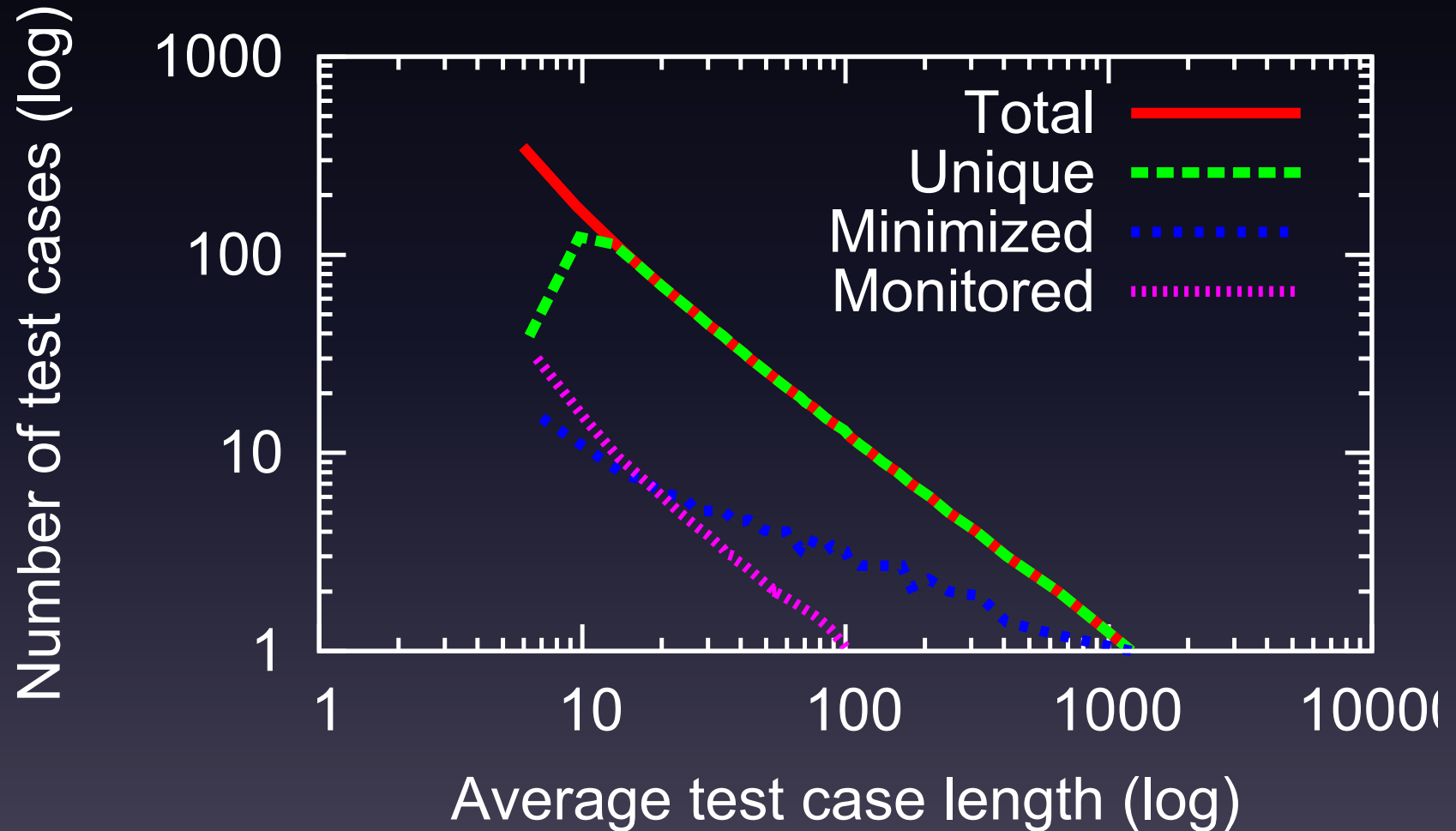
- Are longer tests cheaper?
 - No, if minimization is applied
 - No, for not-quite-so-long test cases
 - Yes, in terms of number of test cases
 - Yes, in terms of total length
- Are longer tests better at detecting faults?
 - A bit.

Conclusions

- Are longer tests cheaper?
 - No, if minimization is applied
 - No, for not-quite-so-long test cases
 - Yes, in terms of number of test cases
 - Yes, in terms of total length
- Are longer tests better at detecting faults?
 - A bit.
- Thank you for your attention!
- Questions?

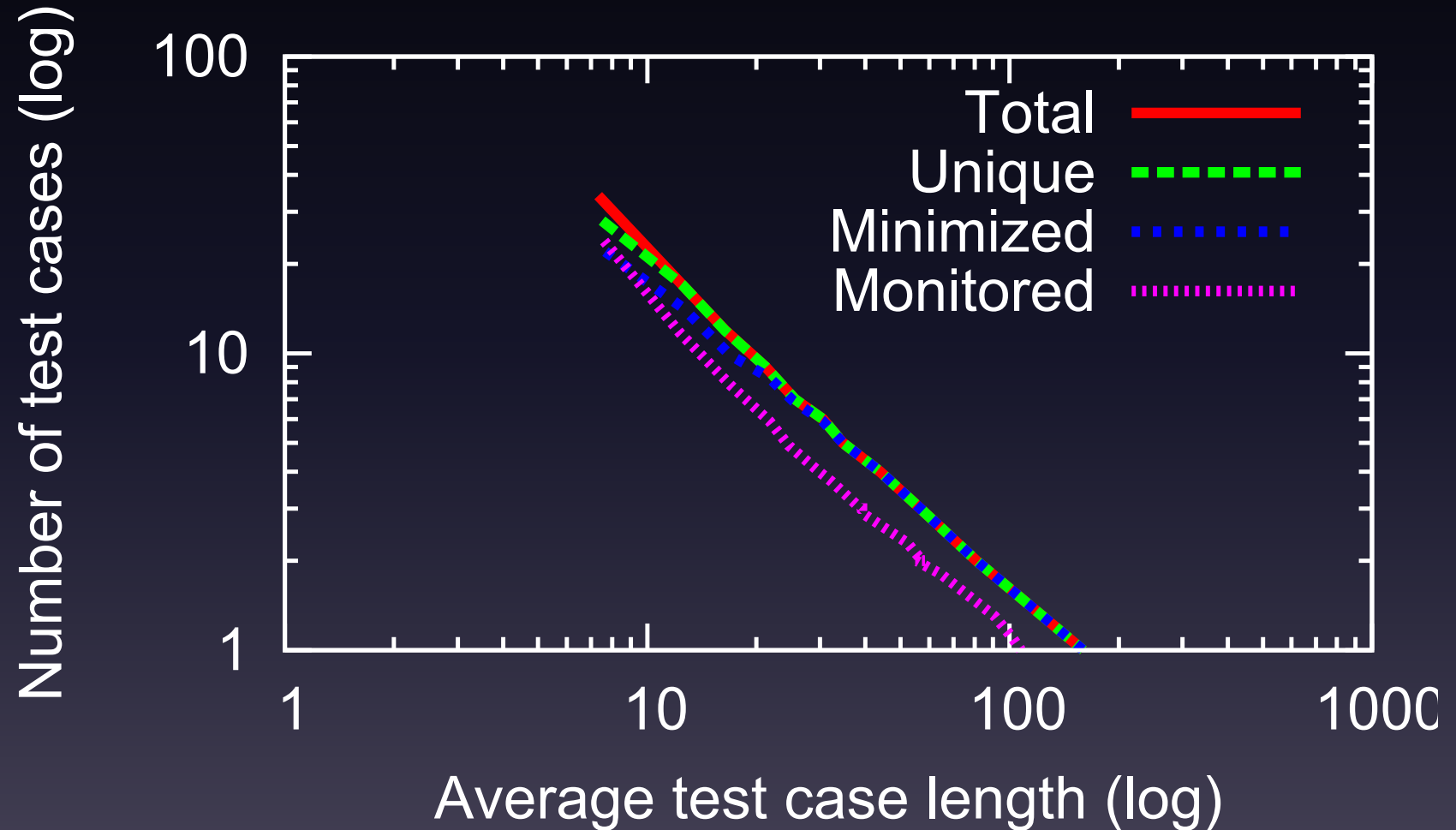
Effects on the Test Suite Size

Cruise, Mutation



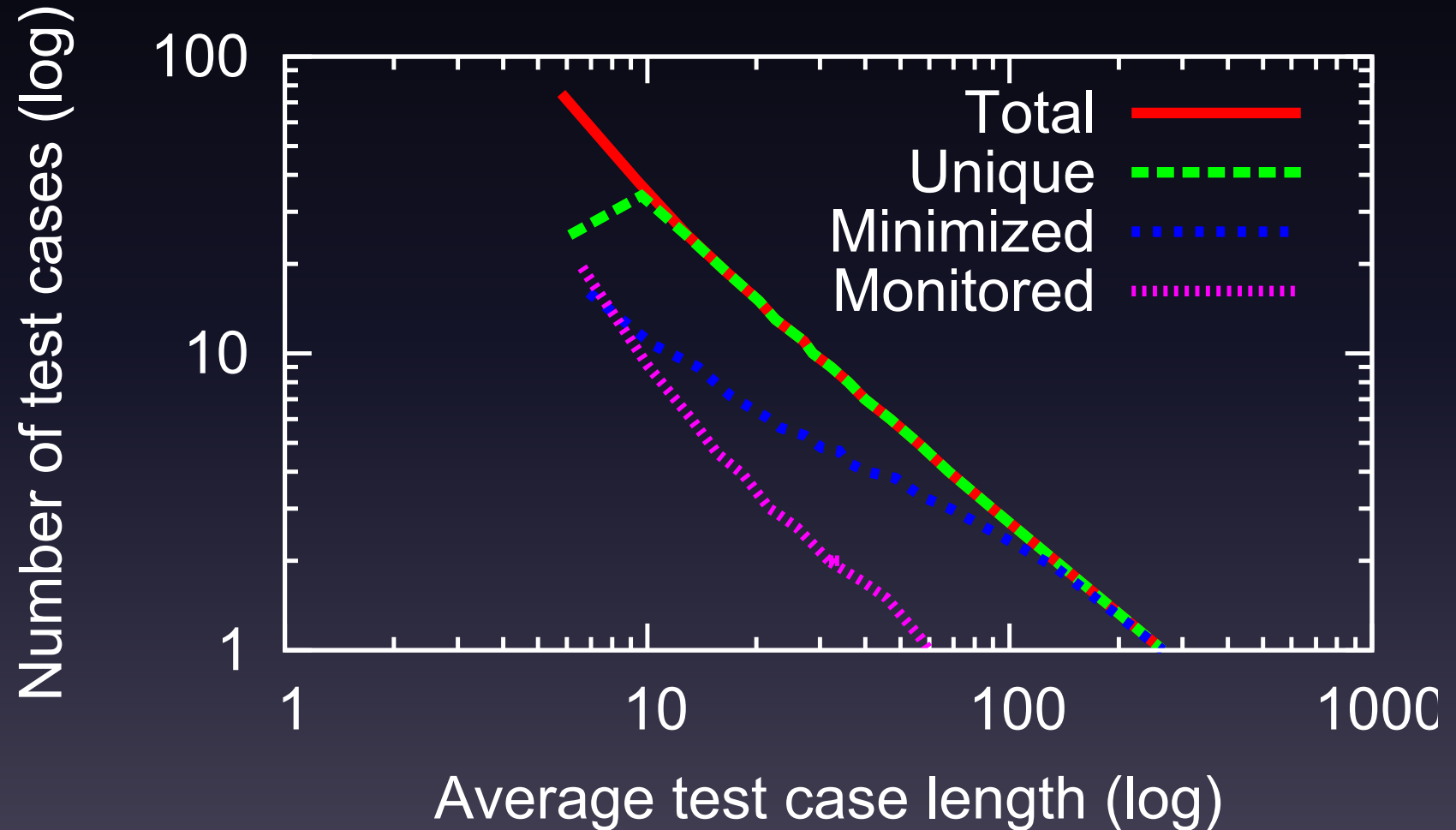
Effects on the Test Suite Size

Cruise, TP



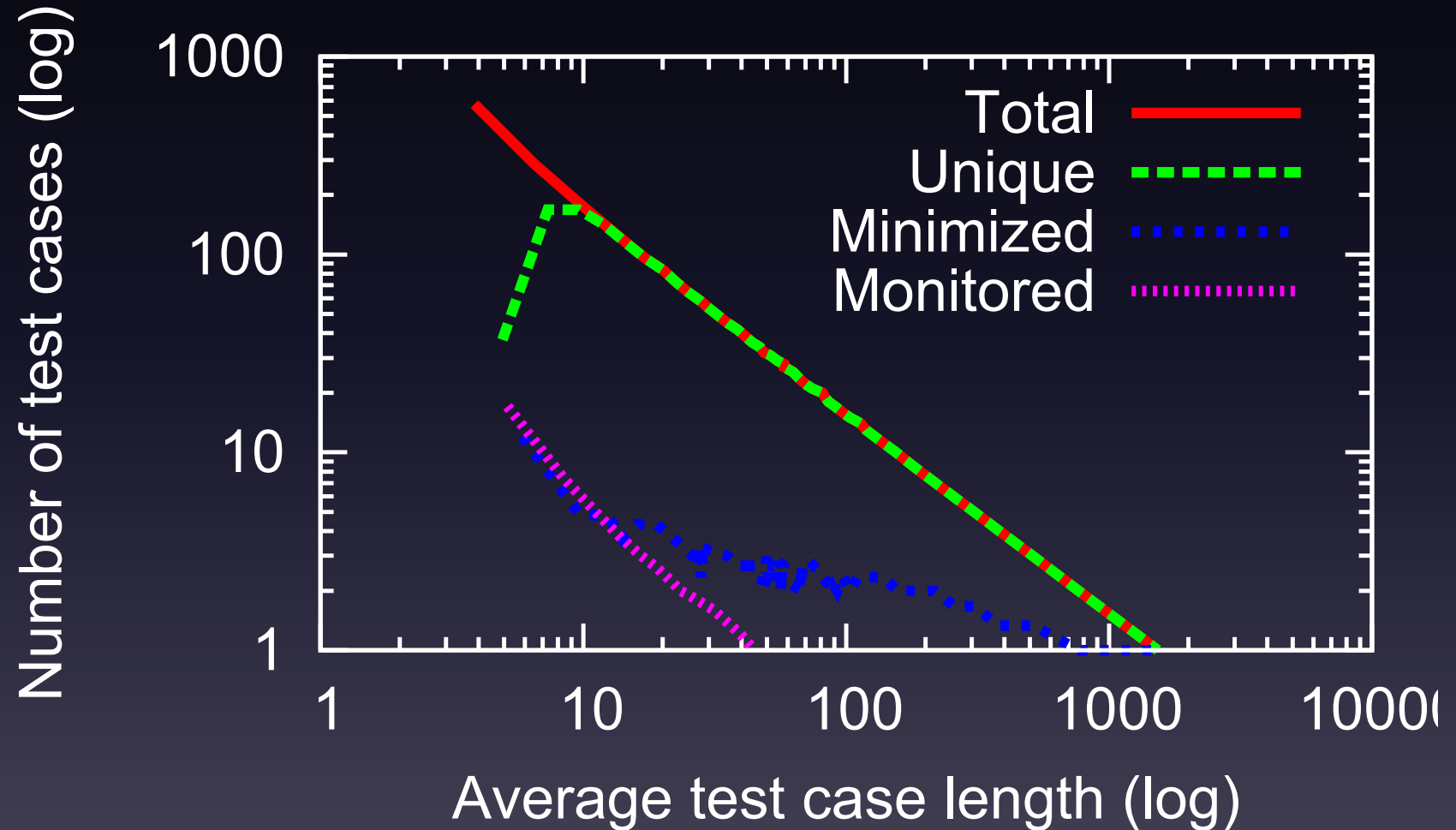
Effects on the Test Suite Size

Cruise, MCDC



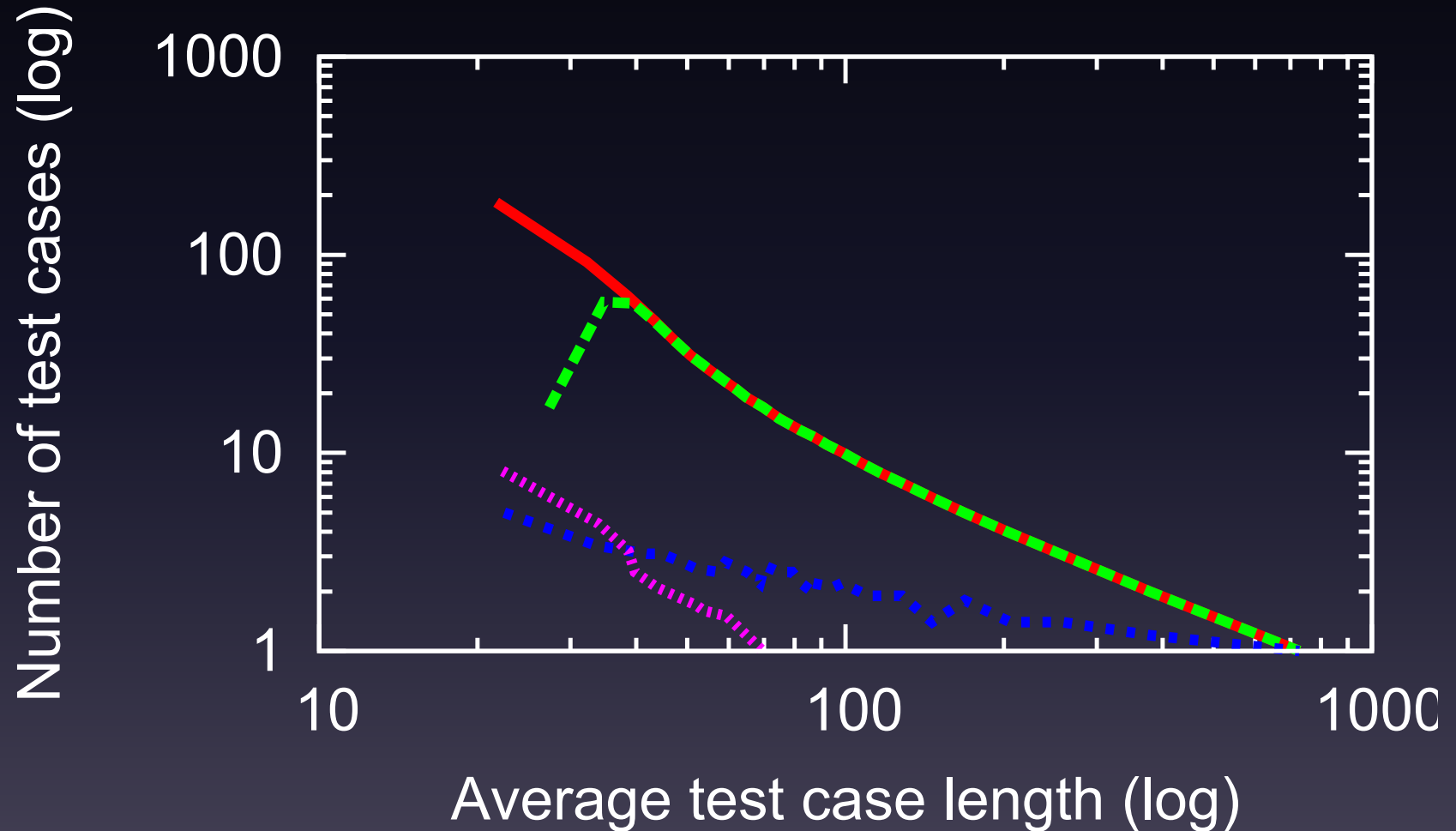
Effects on the Test Suite Size

Cruise, Pairwise



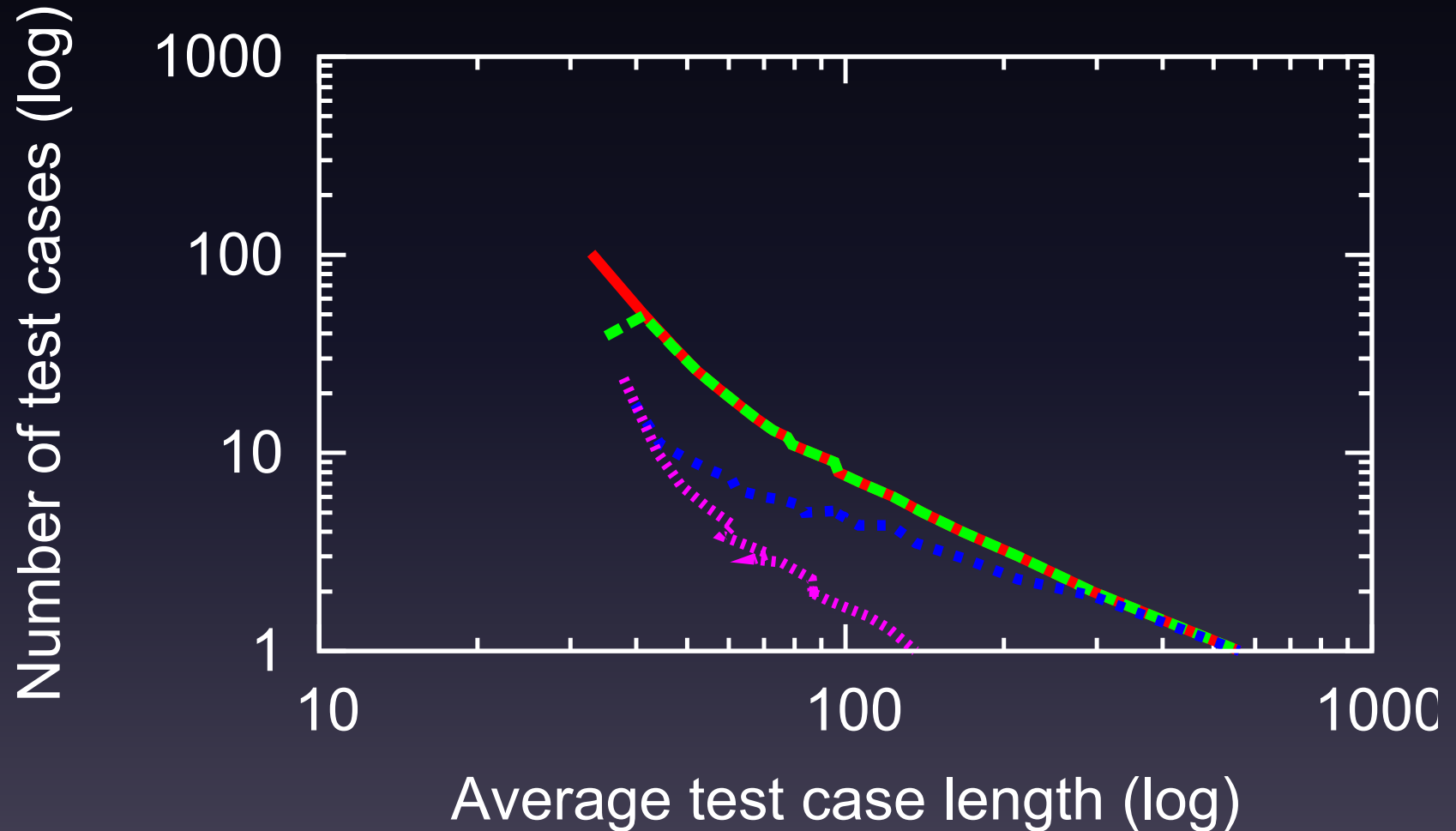
Effects on the Test Suite Size

SIS, Mutation



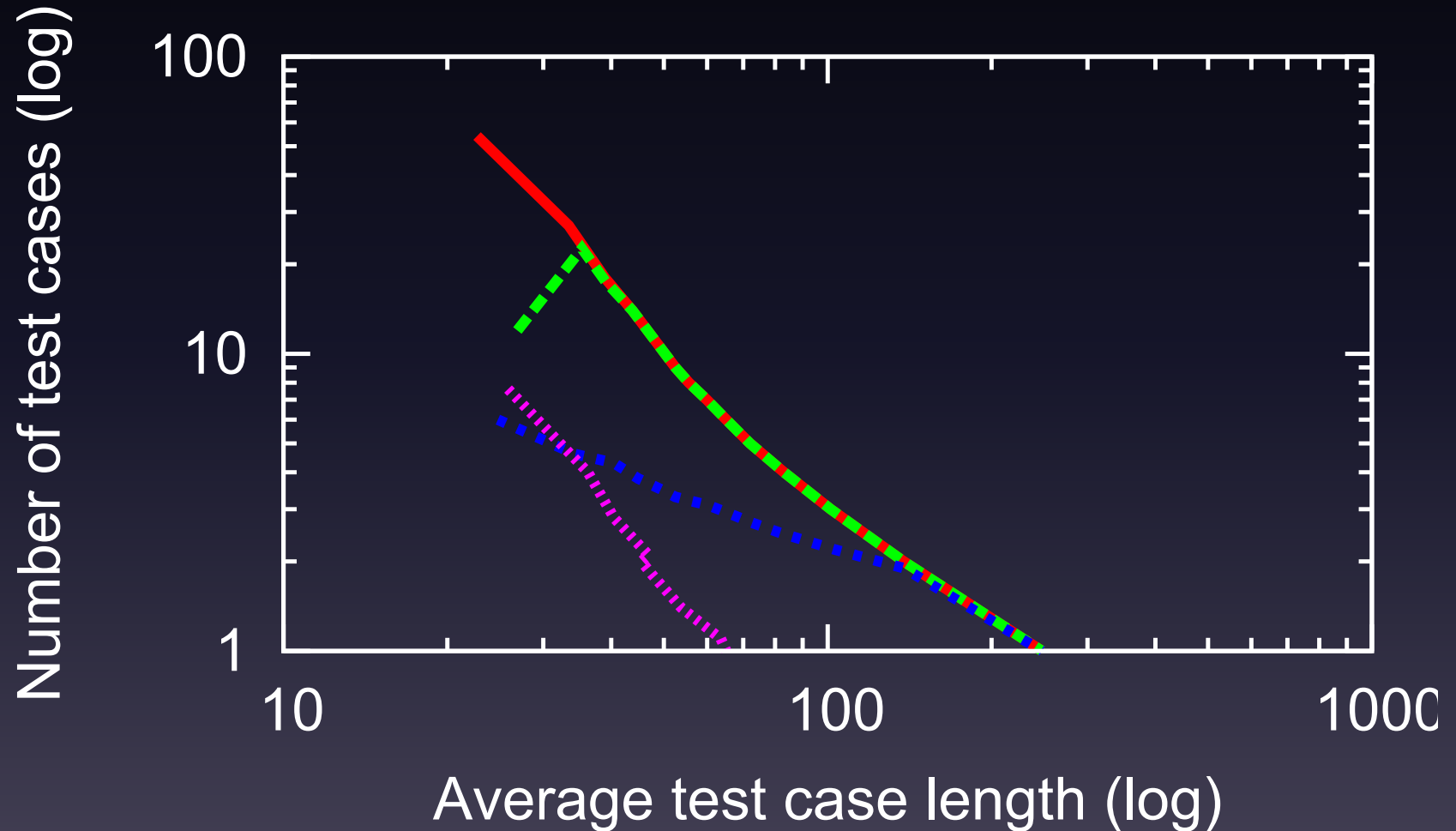
Effects on the Test Suite Size

SIS, TP



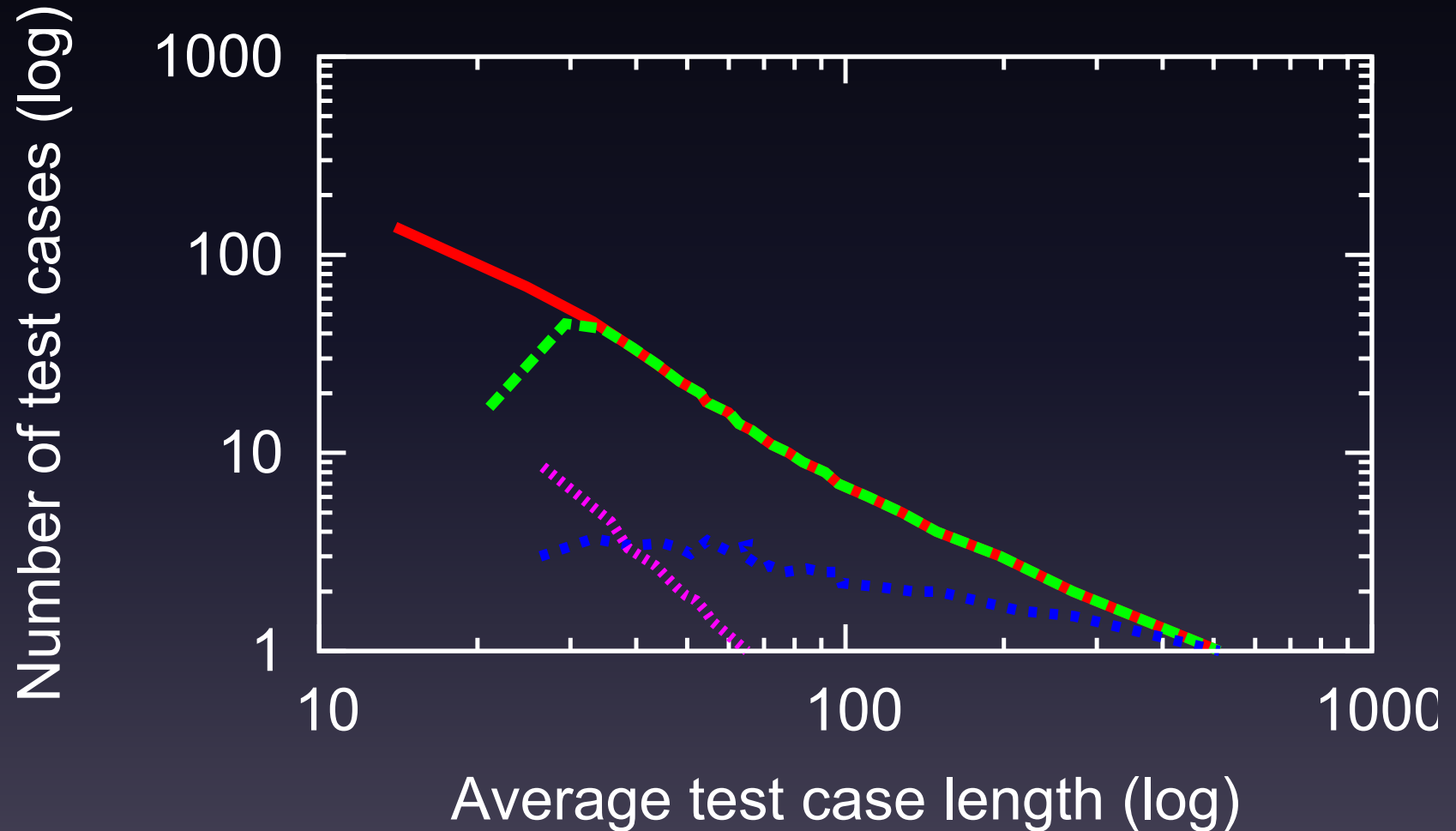
Effects on the Test Suite Size

SIS, MCDC



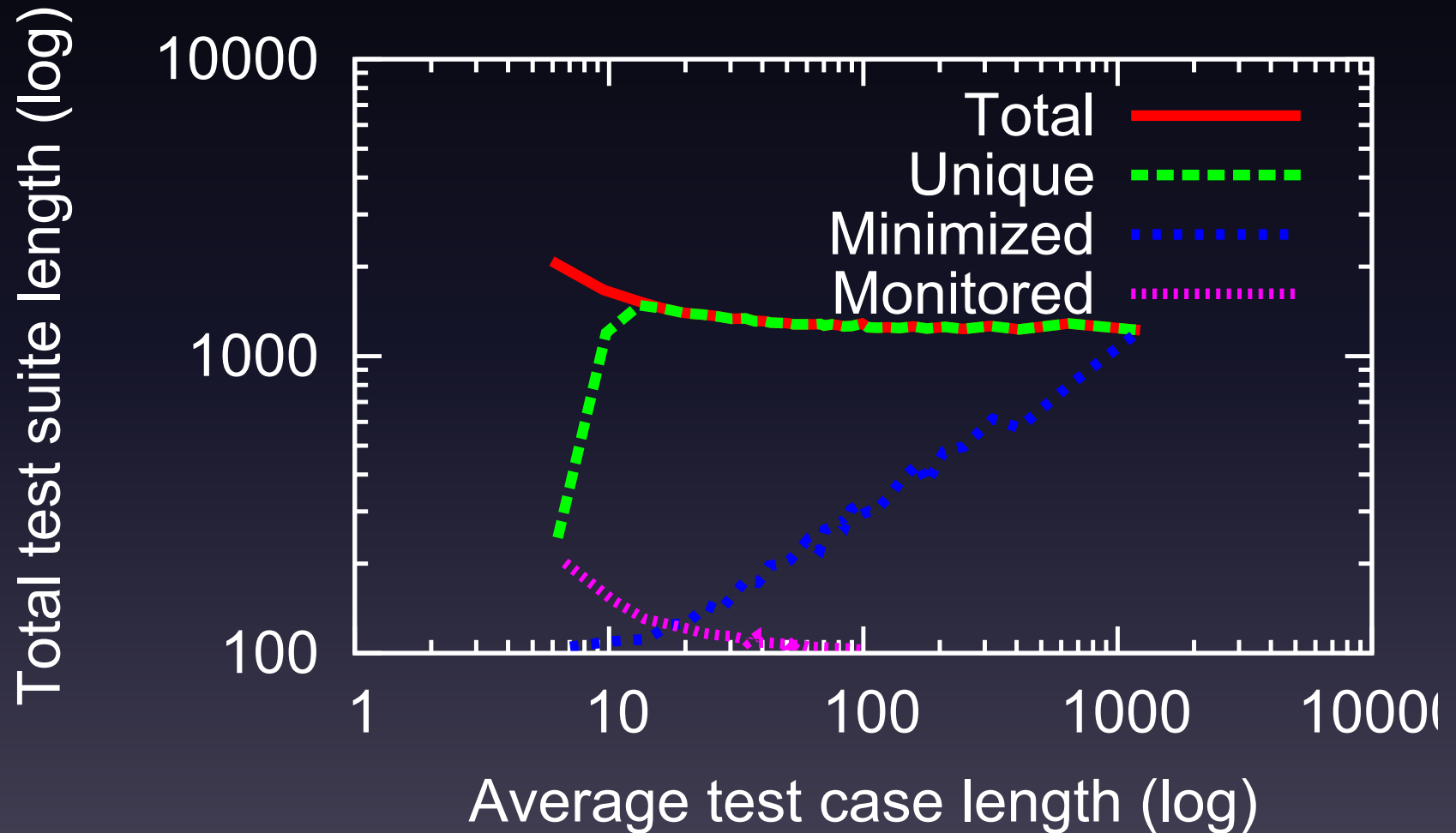
Effects on the Test Suite Size

SIS, Pairwise



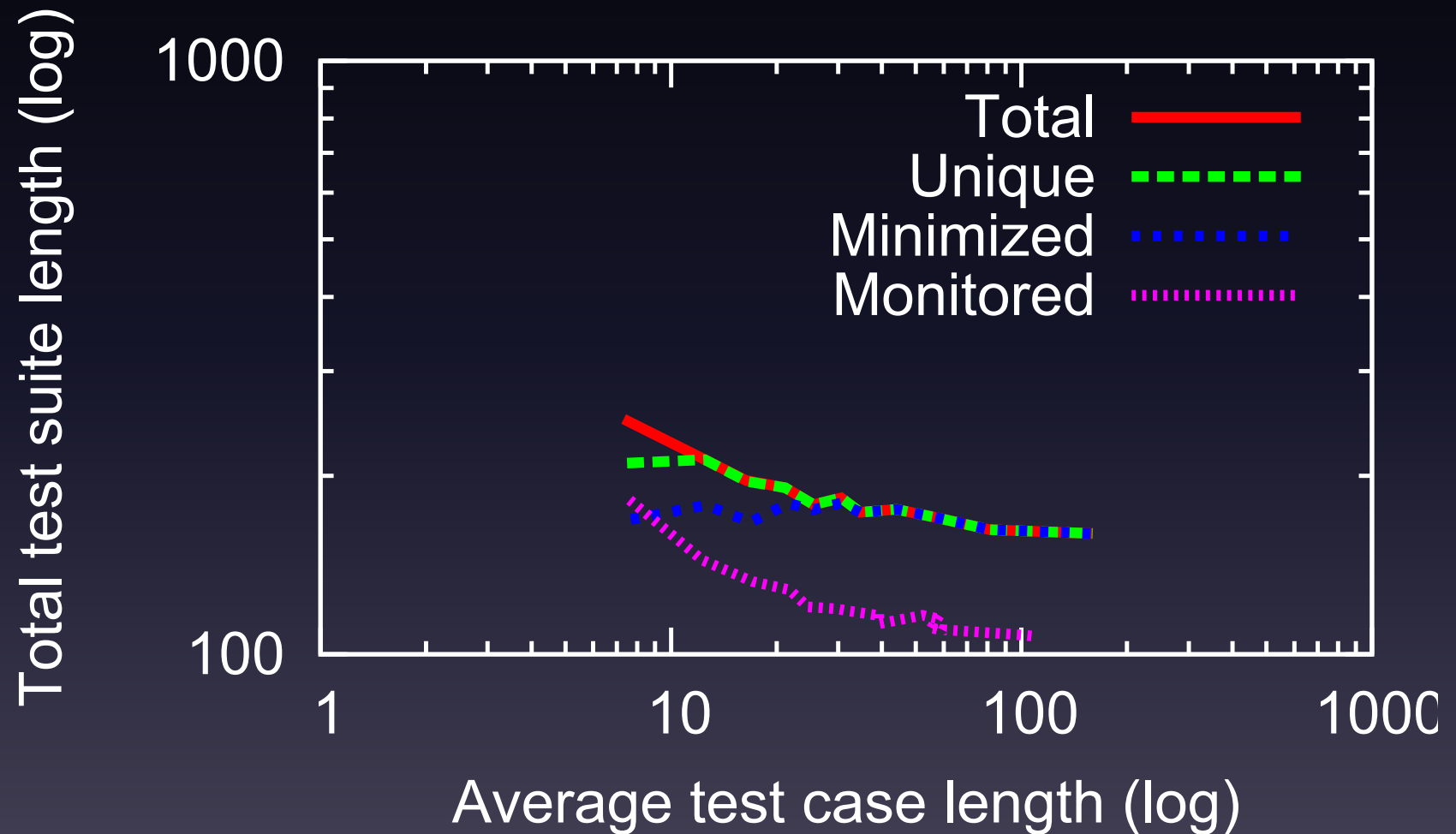
Effects on the Test Suite Length

Cruise, Mutation



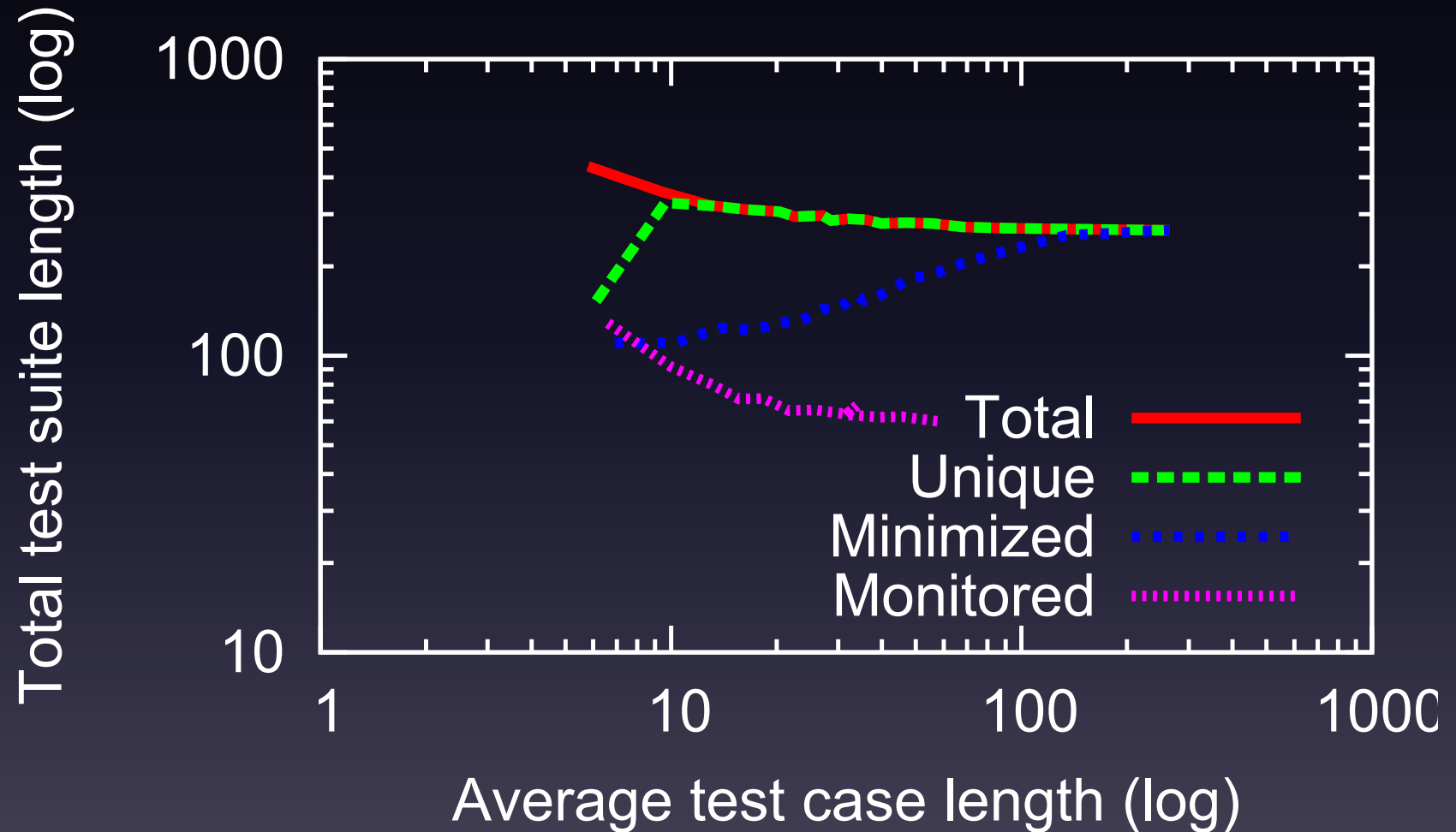
Effects on the Test Suite Length

Cruise, TP



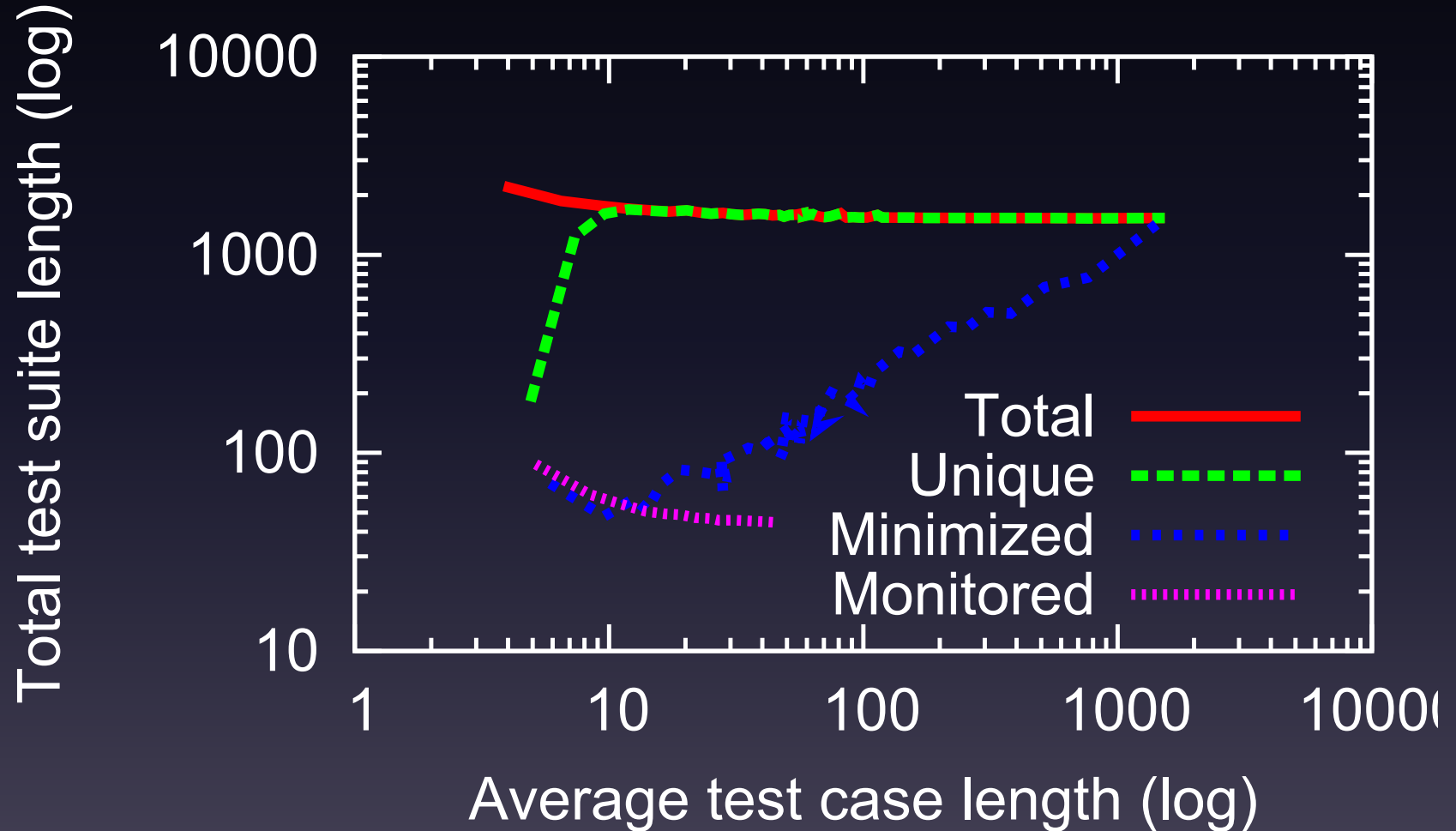
Effects on the Test Suite Length

Cruise, MCDC



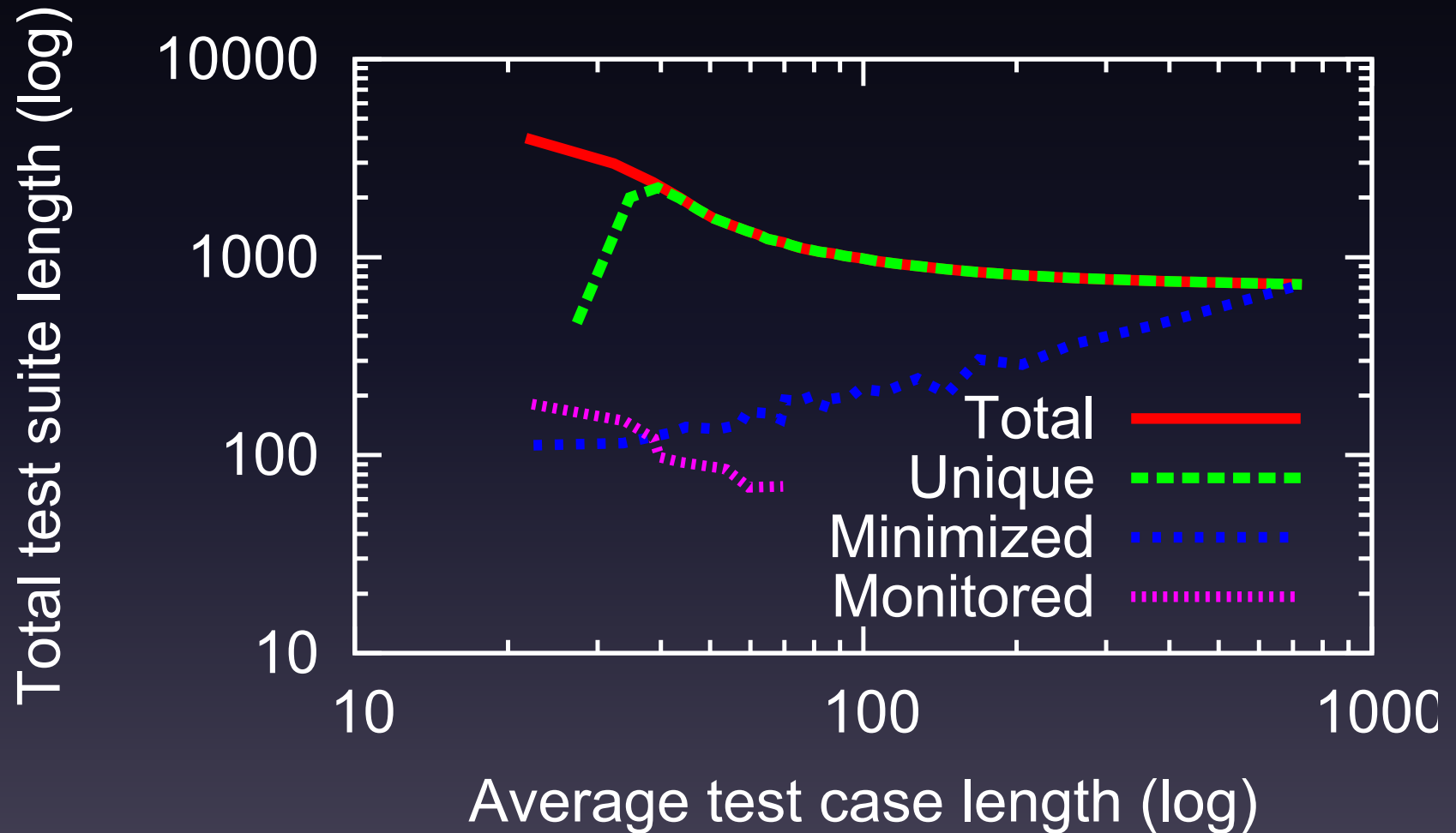
Effects on the Test Suite Length

Cruise, Pairwise



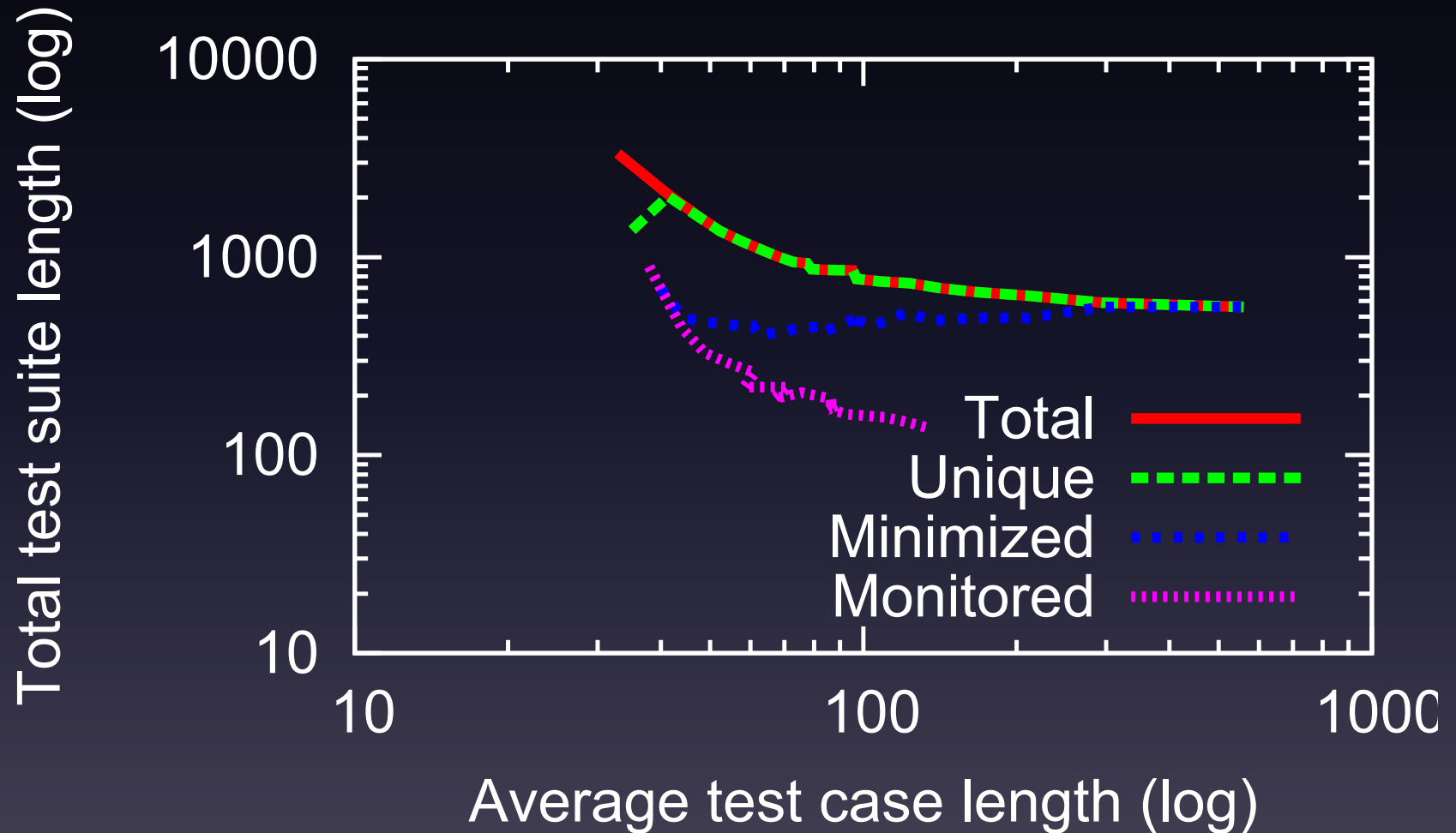
Effects on the Test Suite Length

SIS, Mutation



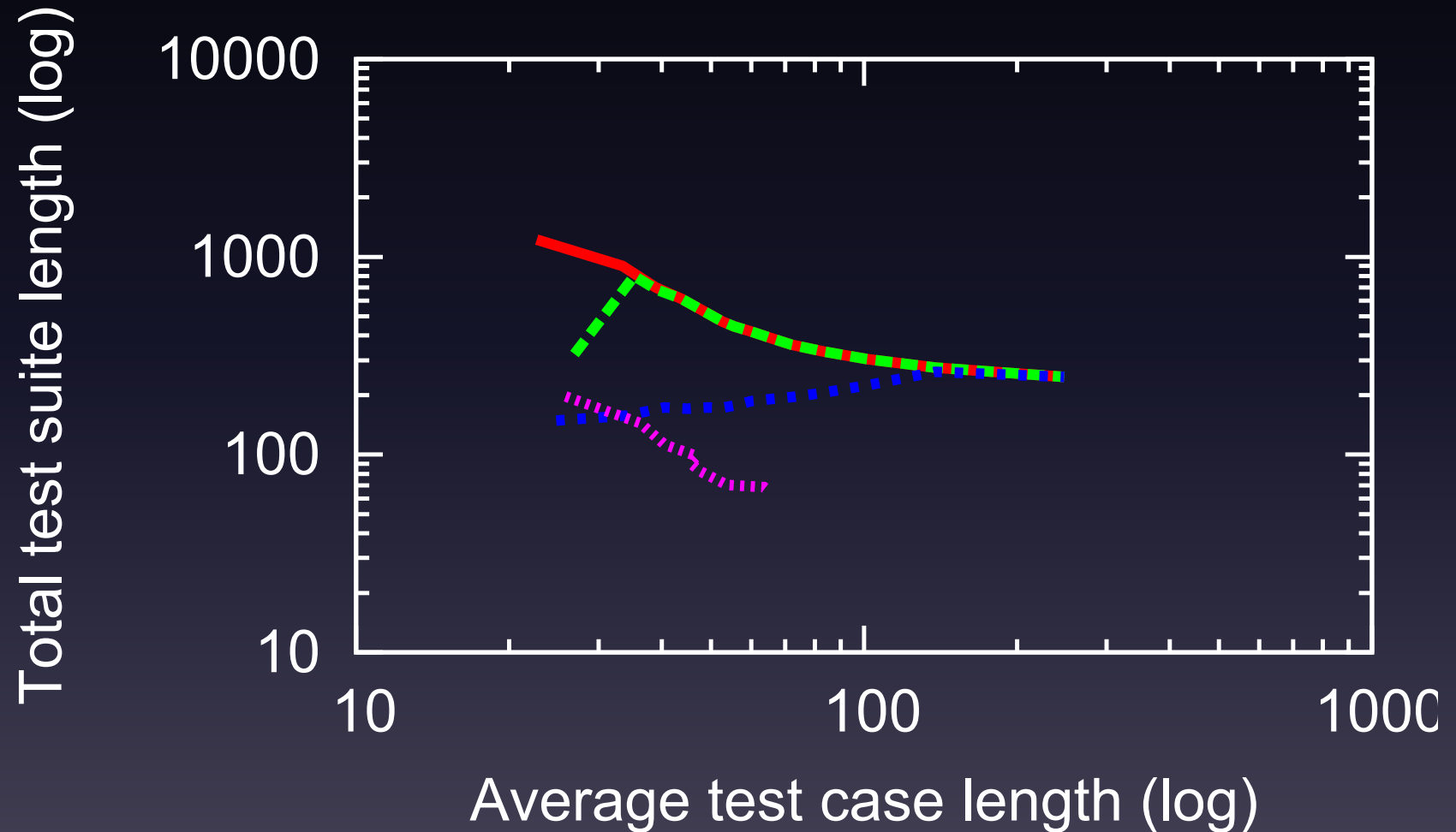
Effects on the Test Suite Length

SIS, TP



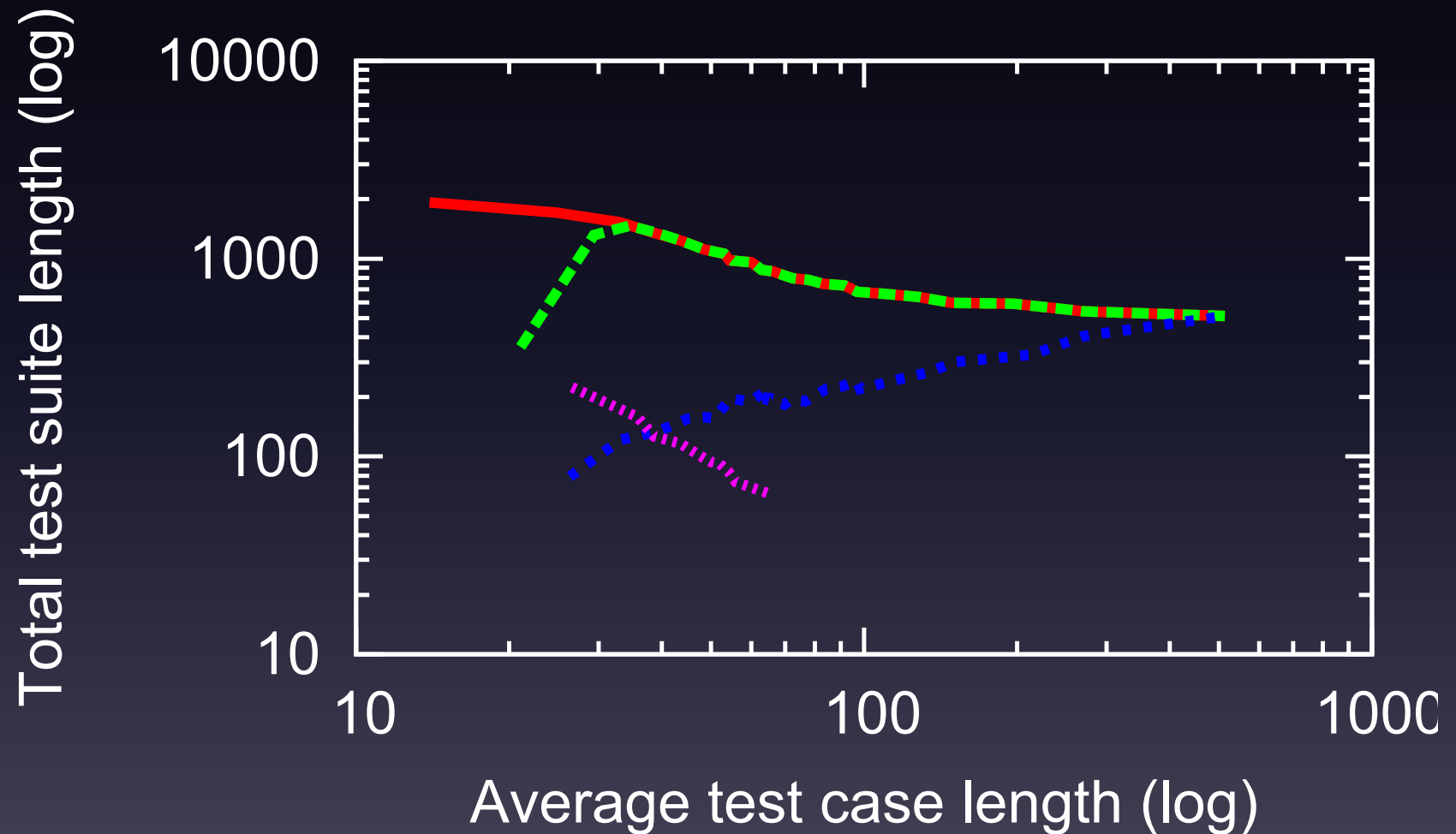
Effects on the Test Suite Length

SIS, MCDC



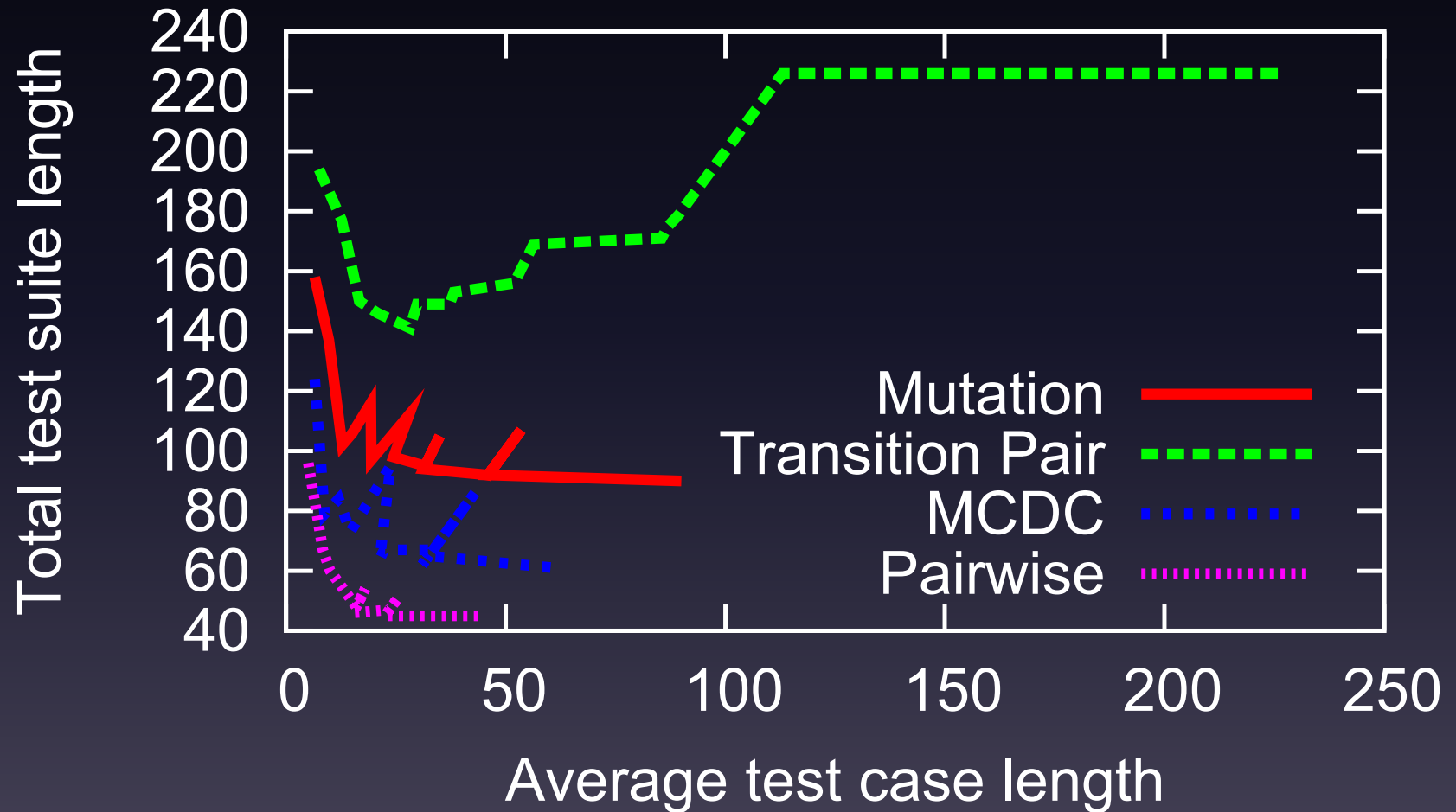
Effects on the Test Suite Length

SIS, Pairwise



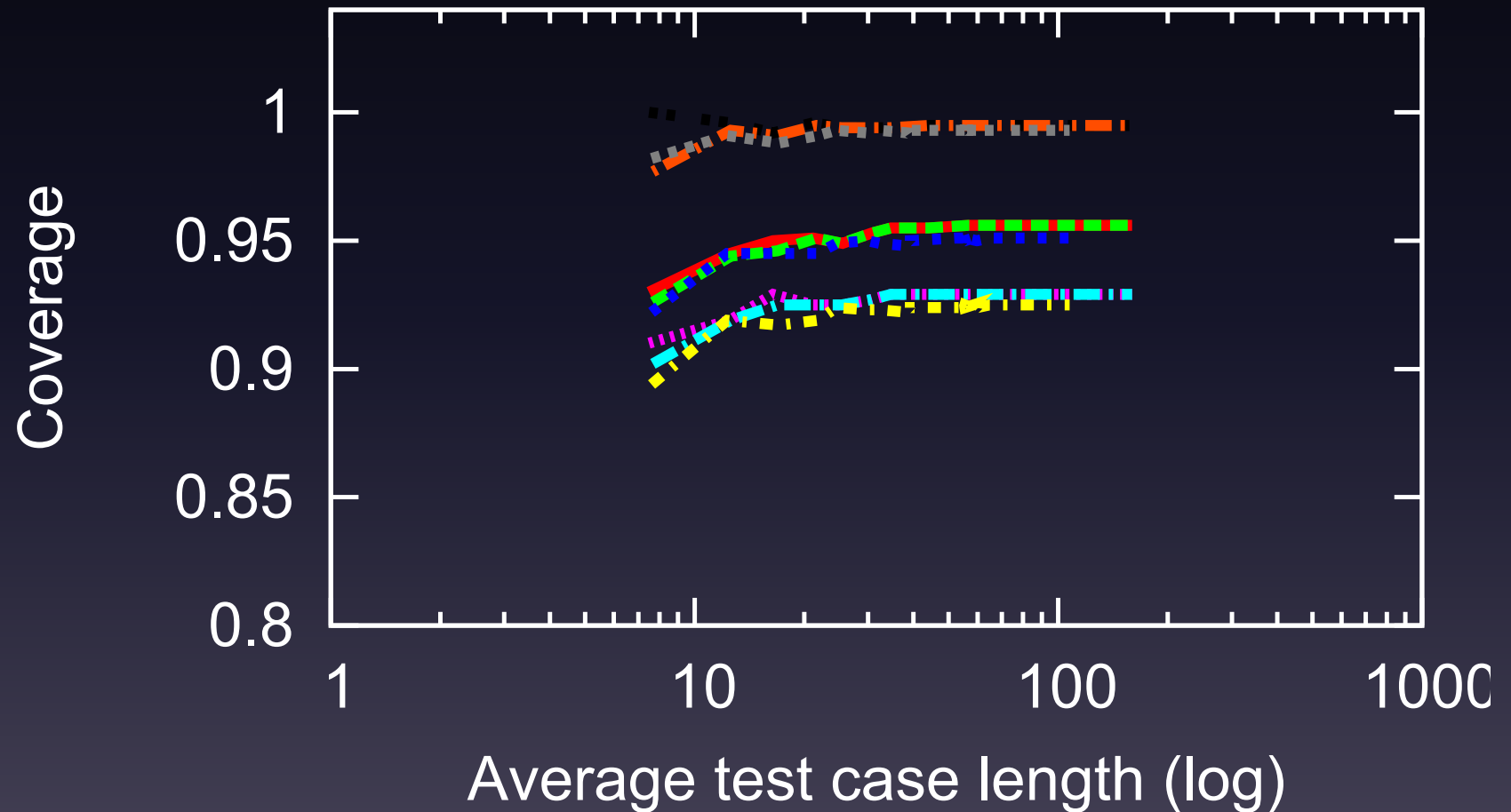
Effects on the Total Length

Cruise, sample run



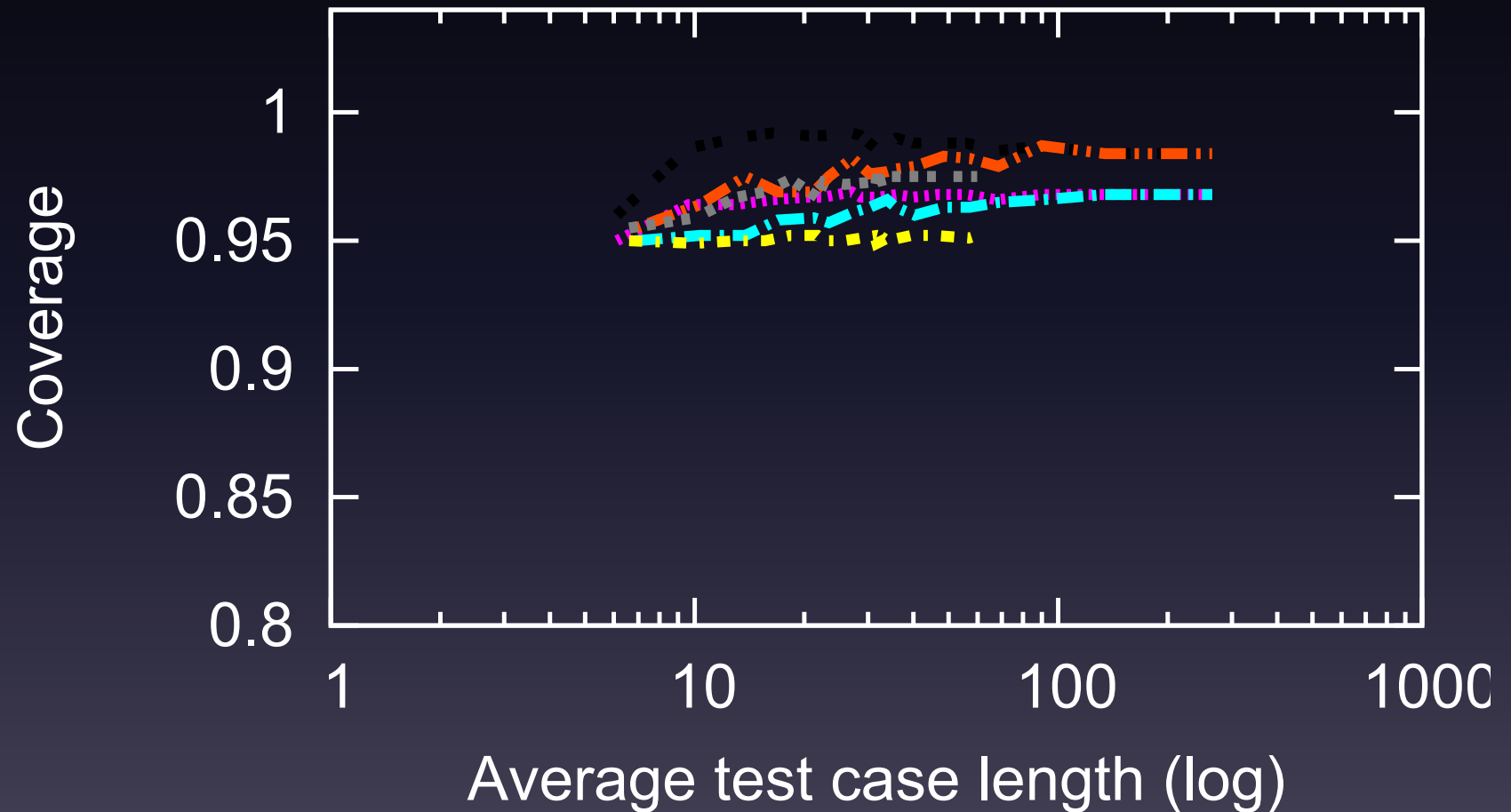
Effects on the Coverage

Cruise, TP



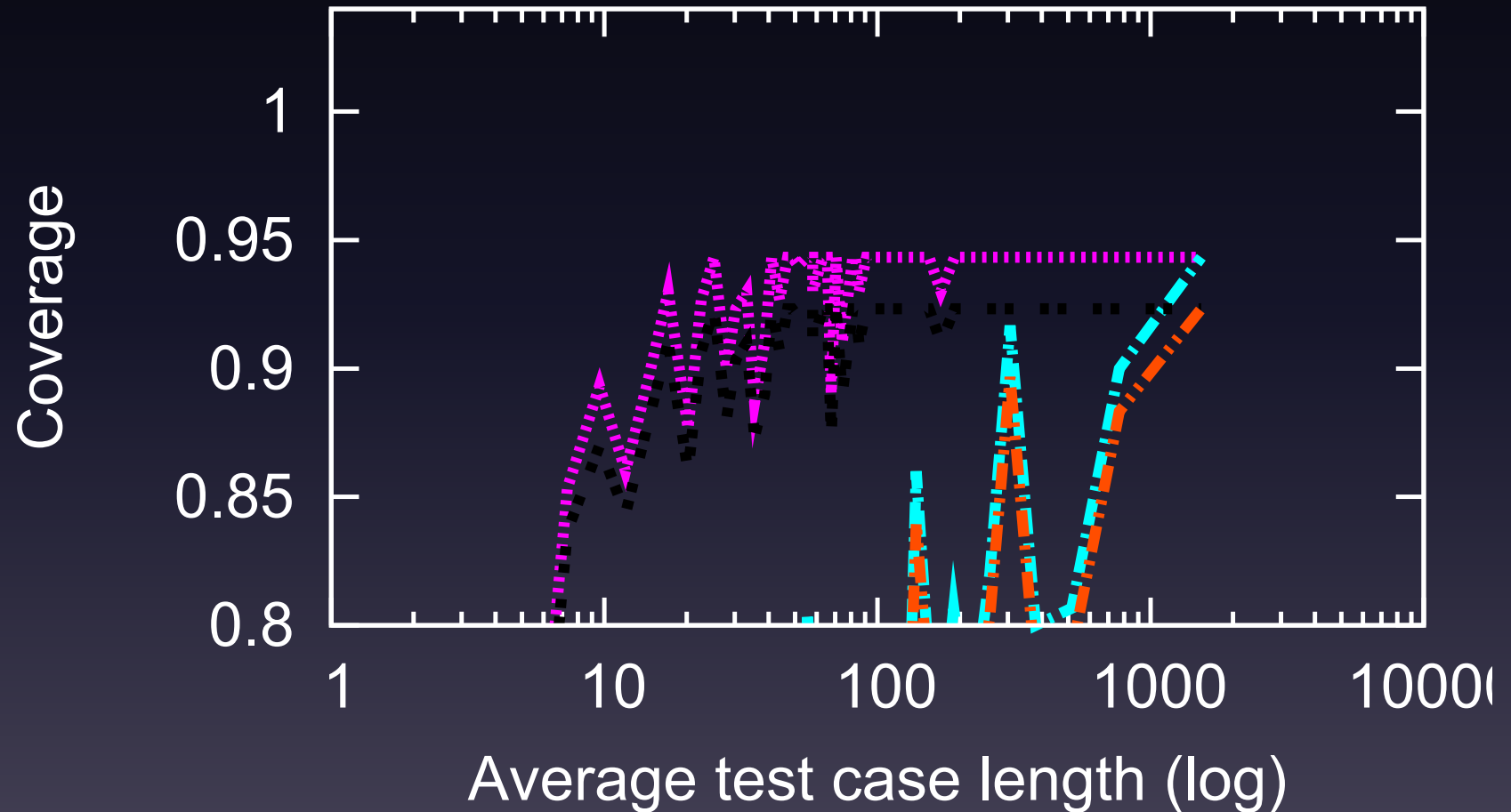
Effects on the Coverage

Cruise, MCDC



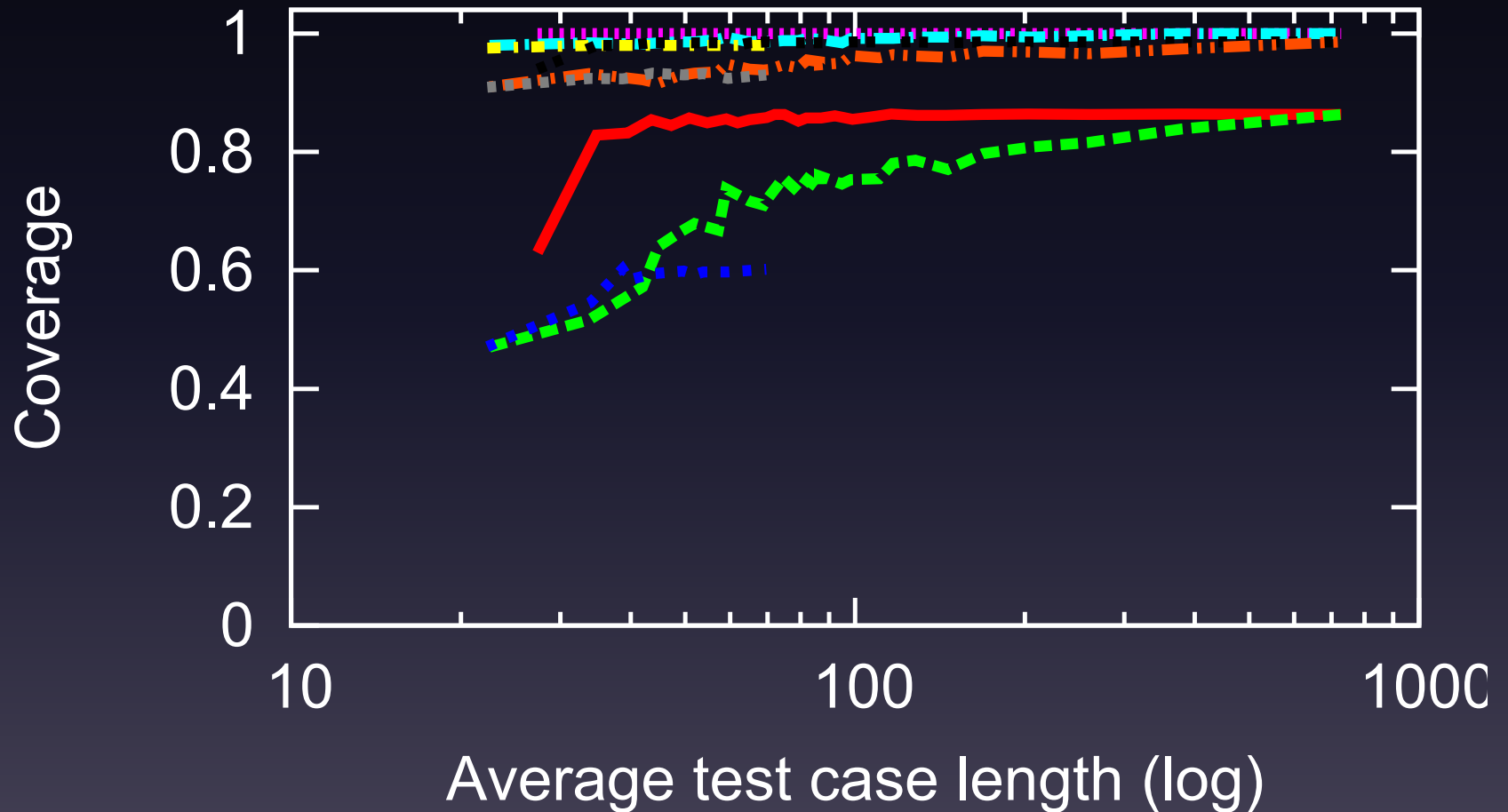
Effects on the Coverage

Cruise, Pairwise



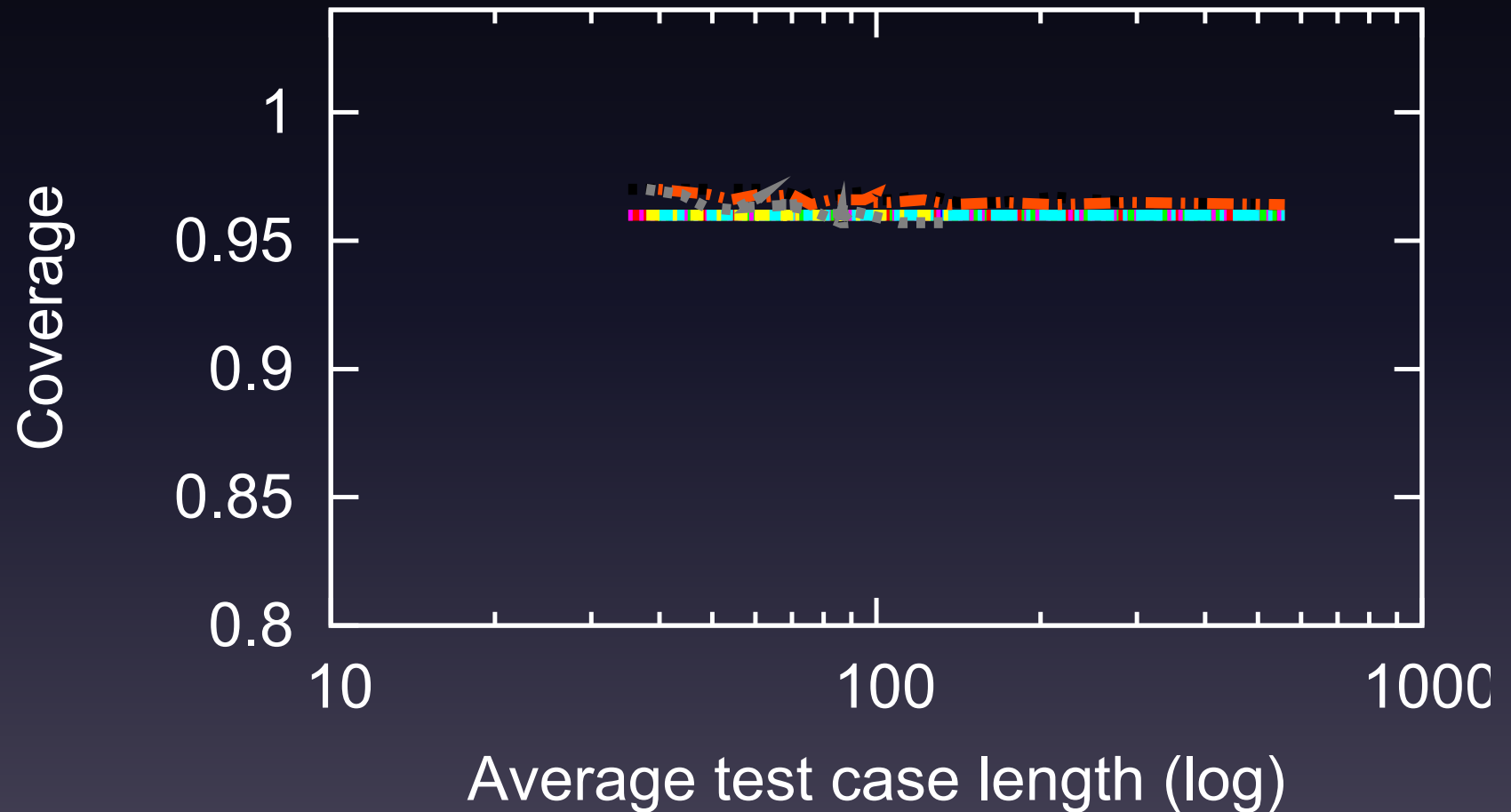
Effects on the Coverage

SIS, Mutation



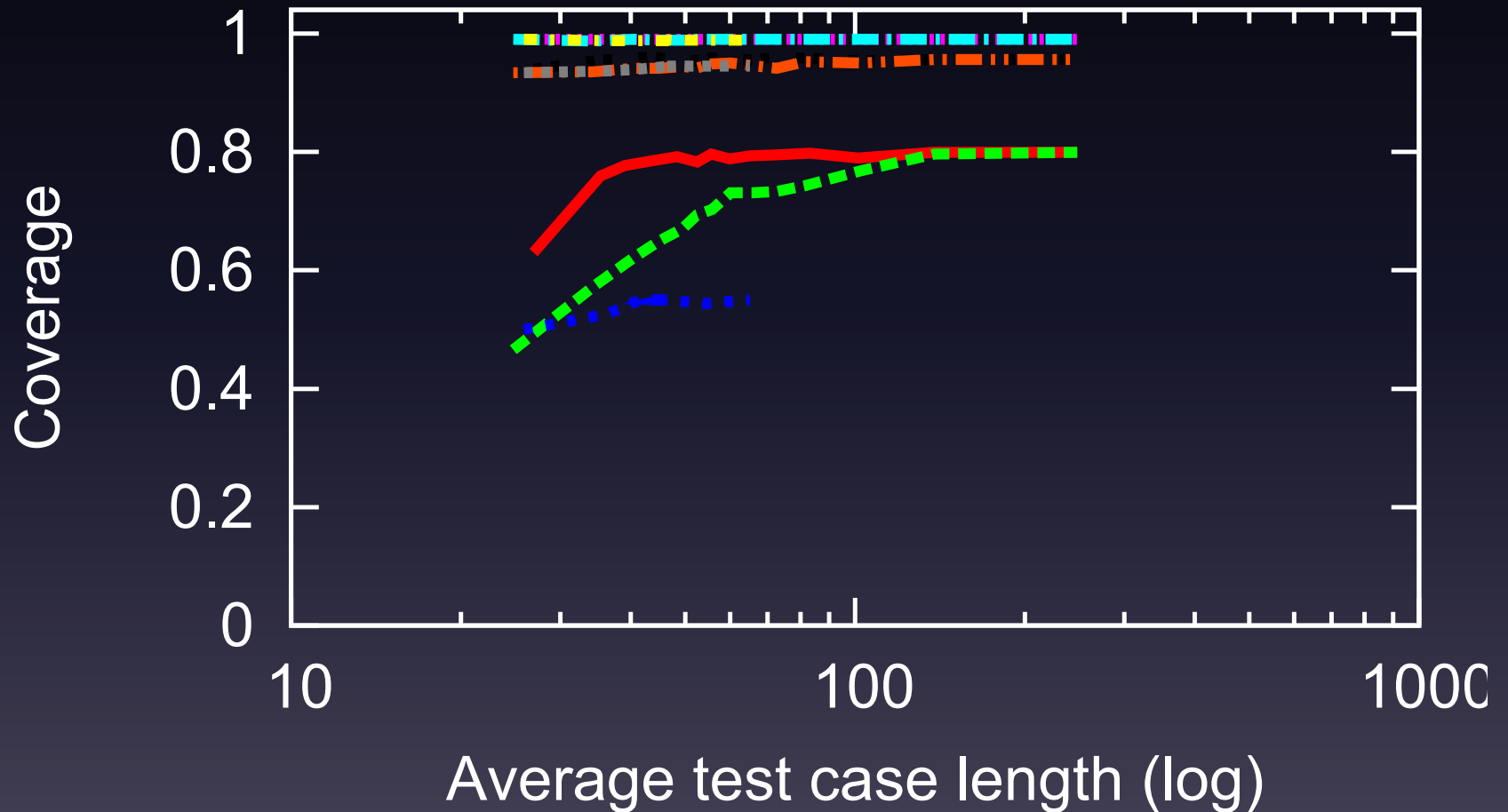
Effects on the Coverage

SIS, TP



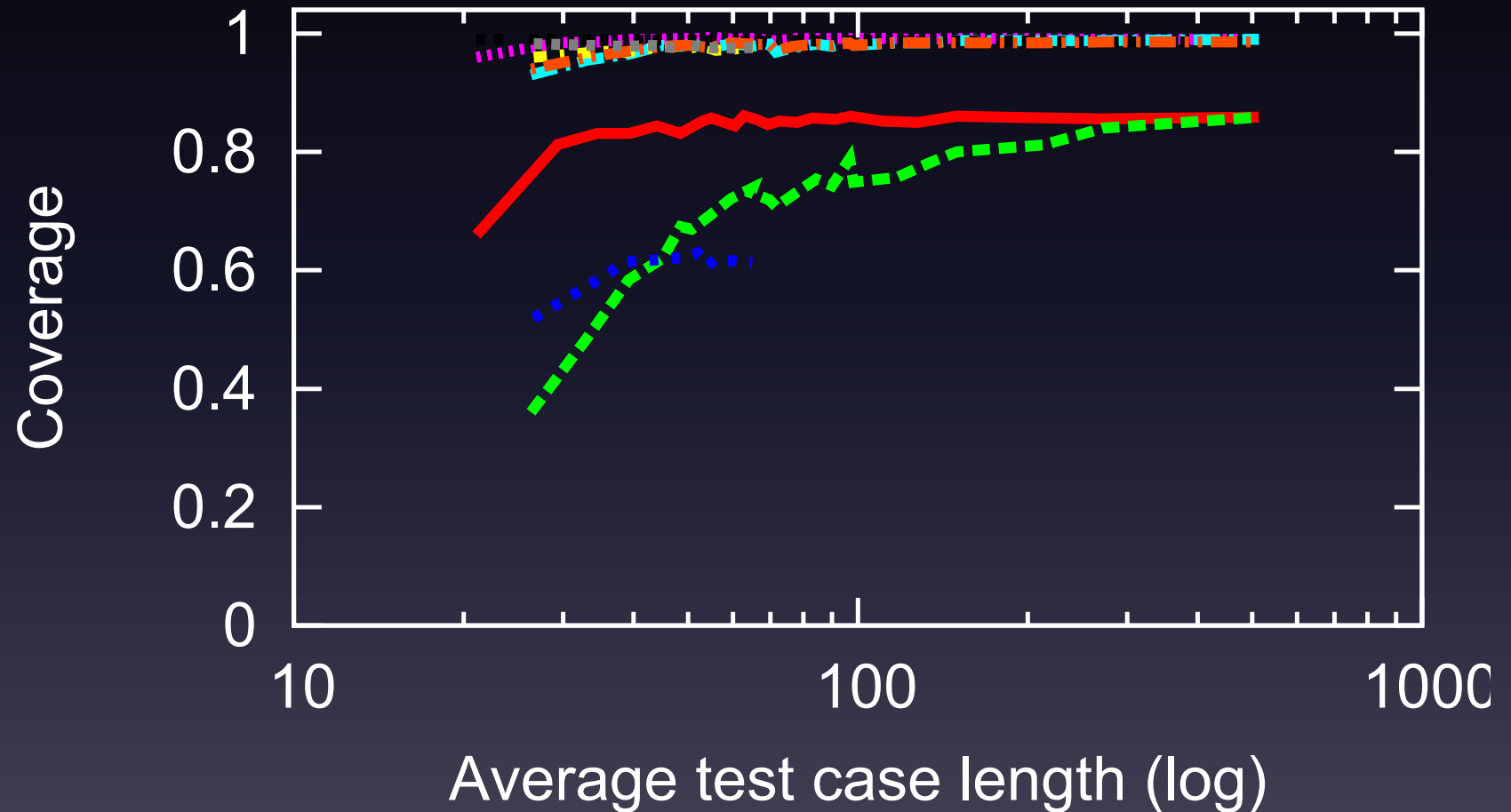
Effects on the Coverage

SIS, MCDC



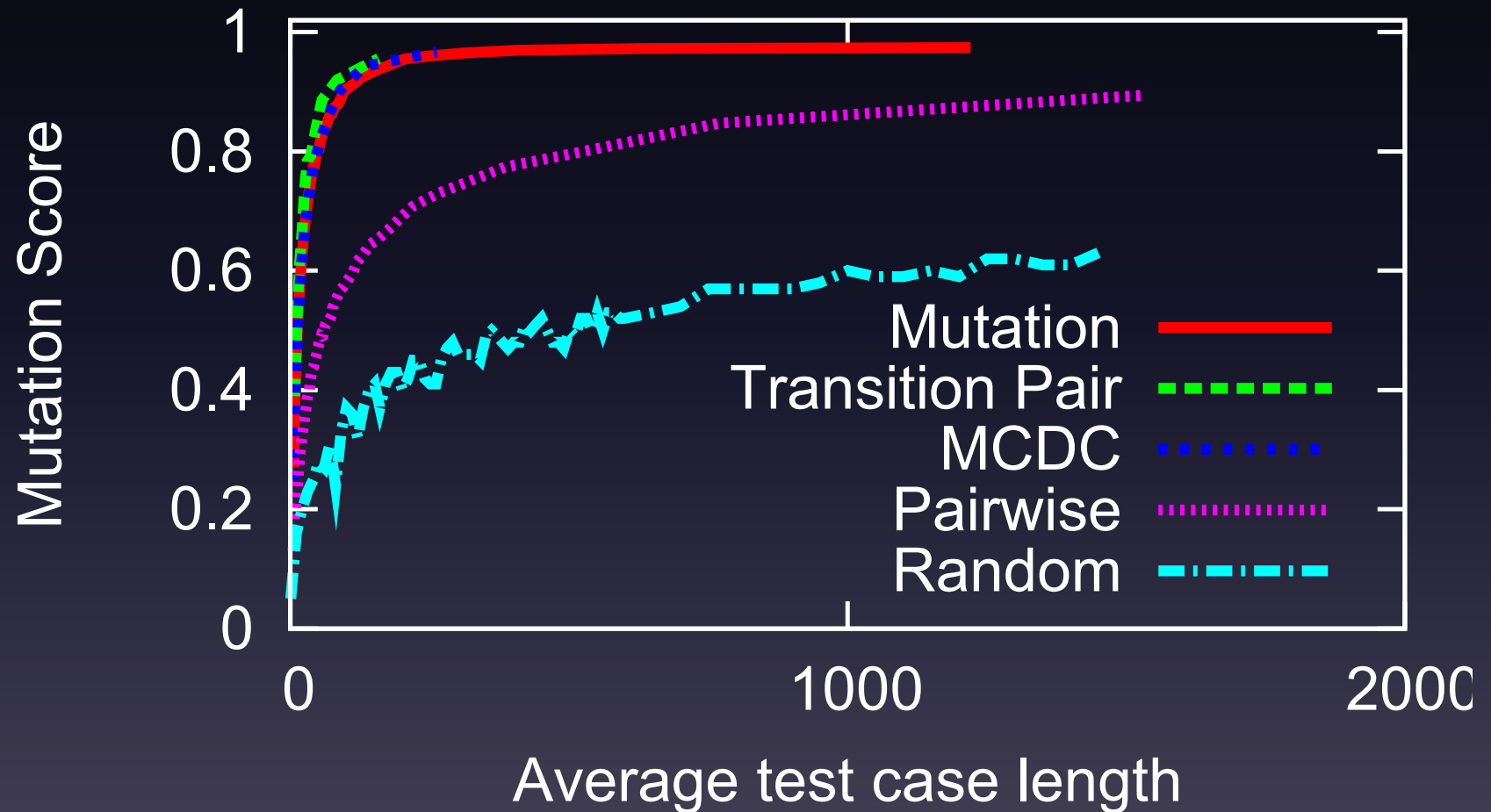
Effects on the Coverage

SIS, Pairwise



Effects on the Mutation Scores

Cruise



Effects on the Mutation Scores

SIS

