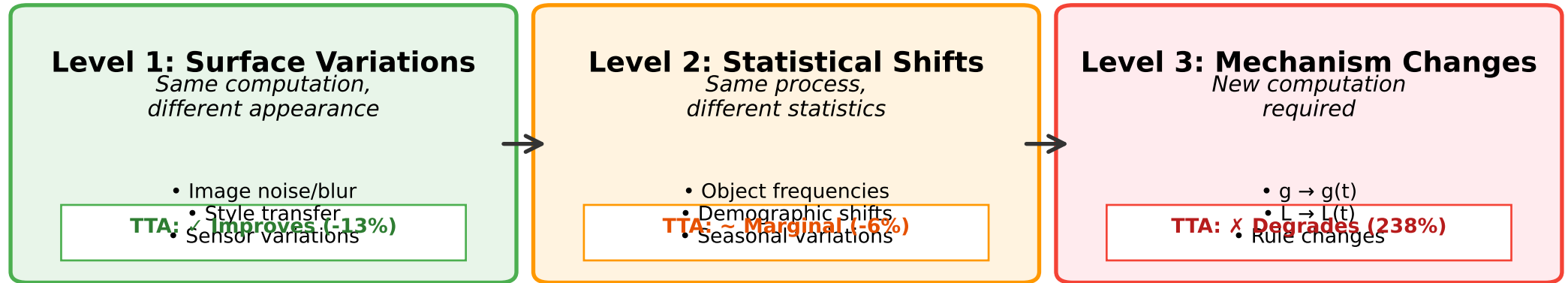


# Distribution Shift Taxonomy



## Physics System Examples

### Two-Ball System

Training:  $g = 9.8$   
Test:  $g(t) = 9.8 + 2\sin(0.1t)$

### Pendulum System

Training:  $L = L_0$   
Test:  $L(t) = L_0(1 + 0.2\sin(0.1t))$

## Key Insight

Test-Time Adaptation (TTA) methods optimize self-supervised objectives that align with true error for Level 1-2 shifts but become misaligned for Level 3 shifts, causing systematic degradation.

**Gradient Alignment:**

Level 1-2:  $\cos(\theta) > 0$

Level 3:  $\cos(\theta) < 0$

**Performance Impact:**

Two-ball: 238% worse

Pendulum: 12-18x worse