



Future Observation



Imagination-
Driven Supervision

*Sec 5.2: World Model as
Imagination Engine*



Robotic Video World Model

Sec 5.1: Core Object



Action
(Action-cond.)

Sec 5.3: Causal Alignment

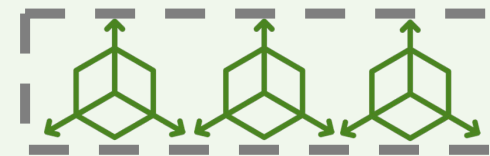


Observation



Language

Instruction
(Lang-cond.)



Structure
(Struct-cond.)

Sec 5.4: Physics Consistent

