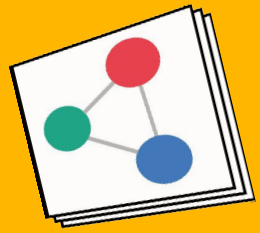
TECHNISCHE UNIVERSITÄT DARMSTADT

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386



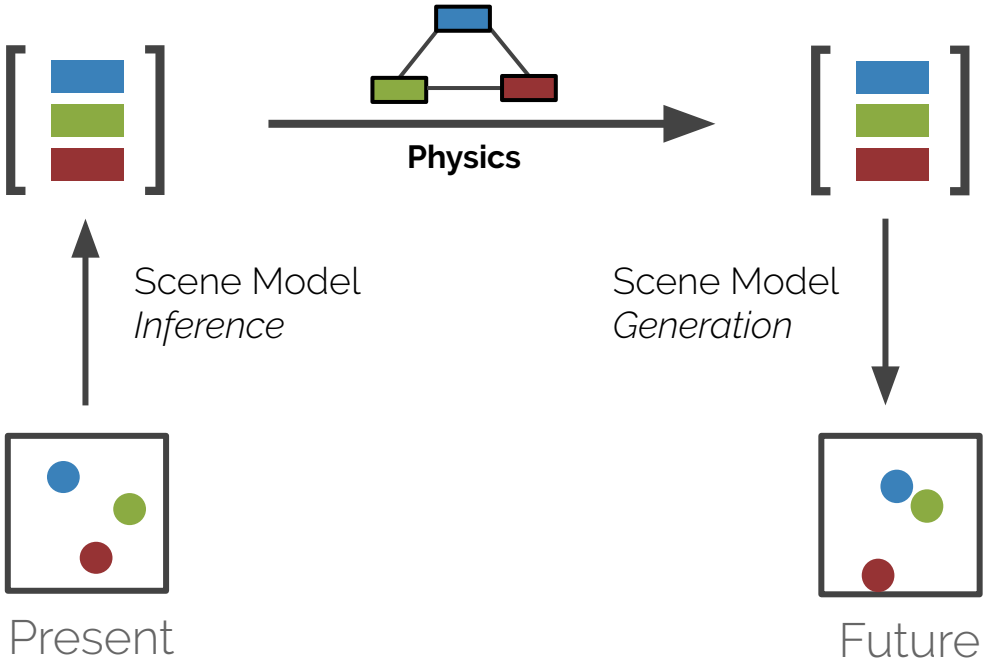
Structured Object-Aware Physics Prediction for Video Modeling and Planning

Jannik Kossen▽*, Karl Stelzner*, Claas Voelcker, Marcel Hussing, Kristian Kersting

(▽ presenting, * equal contribution)

STOVE

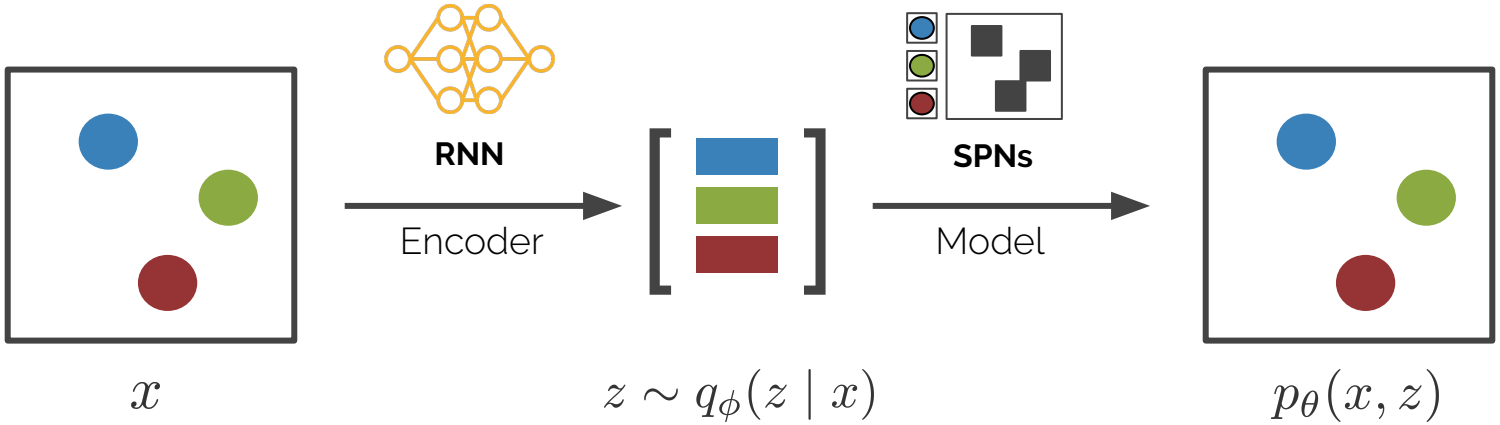
Physical Video Modelling



Kossen et al., Structured Object-Aware Physics Prediction for Video Modeling and Planning

SuPAIR – Scene Modeling

Sum-Product Attend-Infer-Repeat

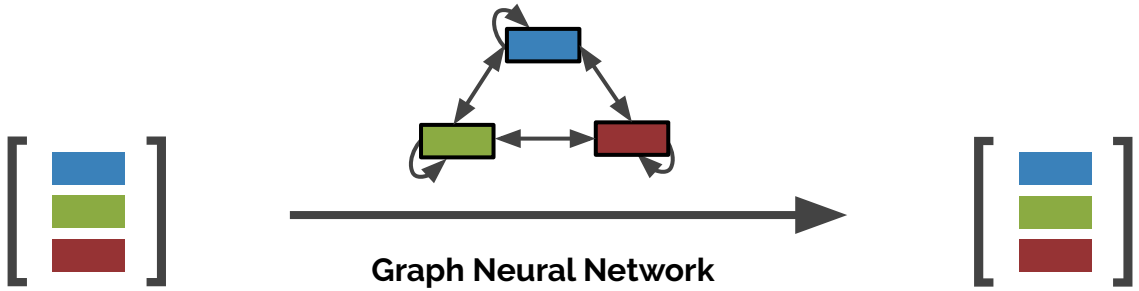


(Stelzner et al., 2019)

Kossen et al., Structured Object-Aware Physics Prediction for Video Modeling and Planning

Relational Physics Prediction

Graph Neural Network



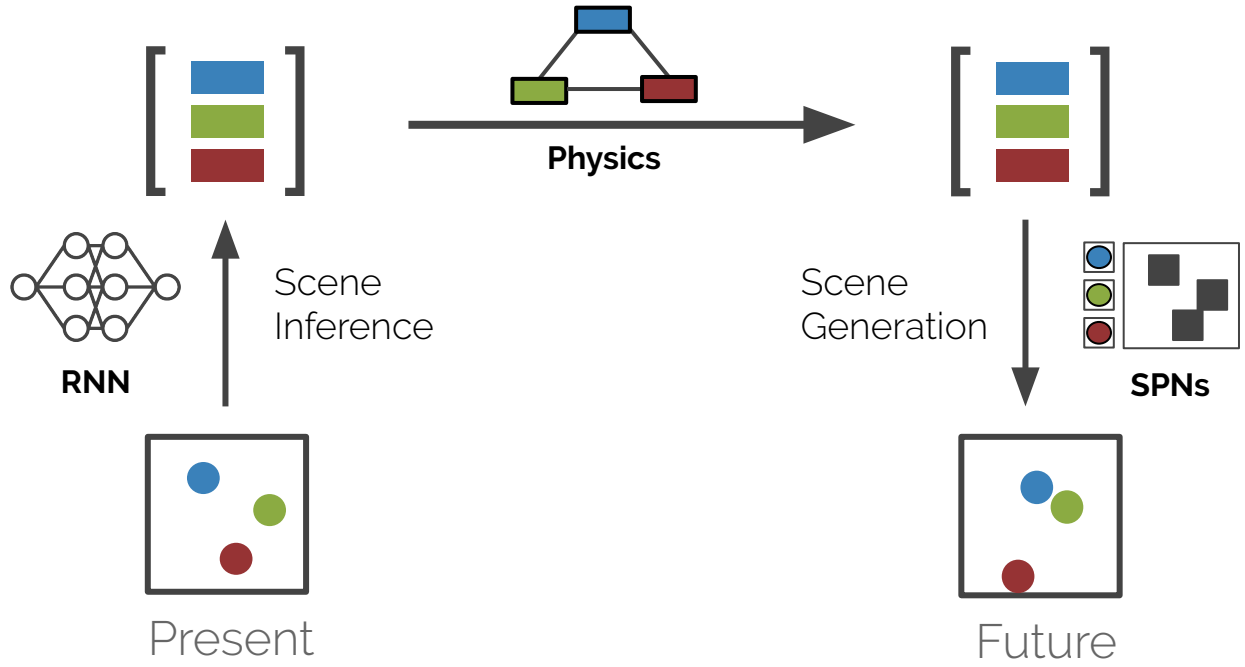
$$z_{t-1}$$

$$z_t \sim p(z_t \mid z_{t-1})$$

(NPE – Chang et. al, 2017; VIN – Watters et al., 2019; R-NEM – Van Steenkiste et al. 2018)

STOVE – The Model

Structured, Object-Aware, Physics

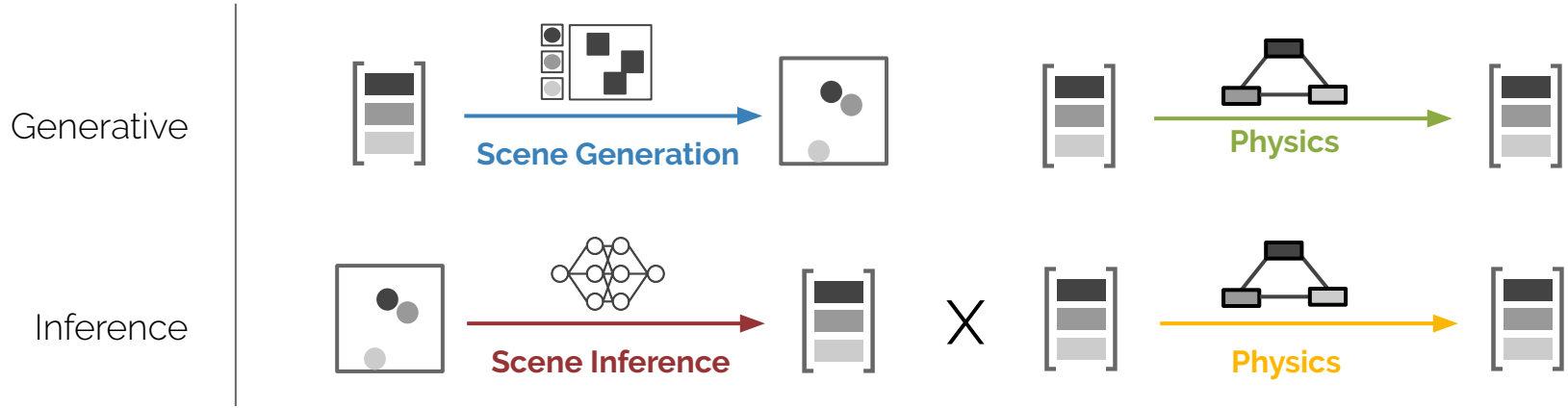


Kossen et al., Structured Object-Aware Physics Prediction for Video Modeling and Planning

The *STOVE* Model

Maximising the ELBO

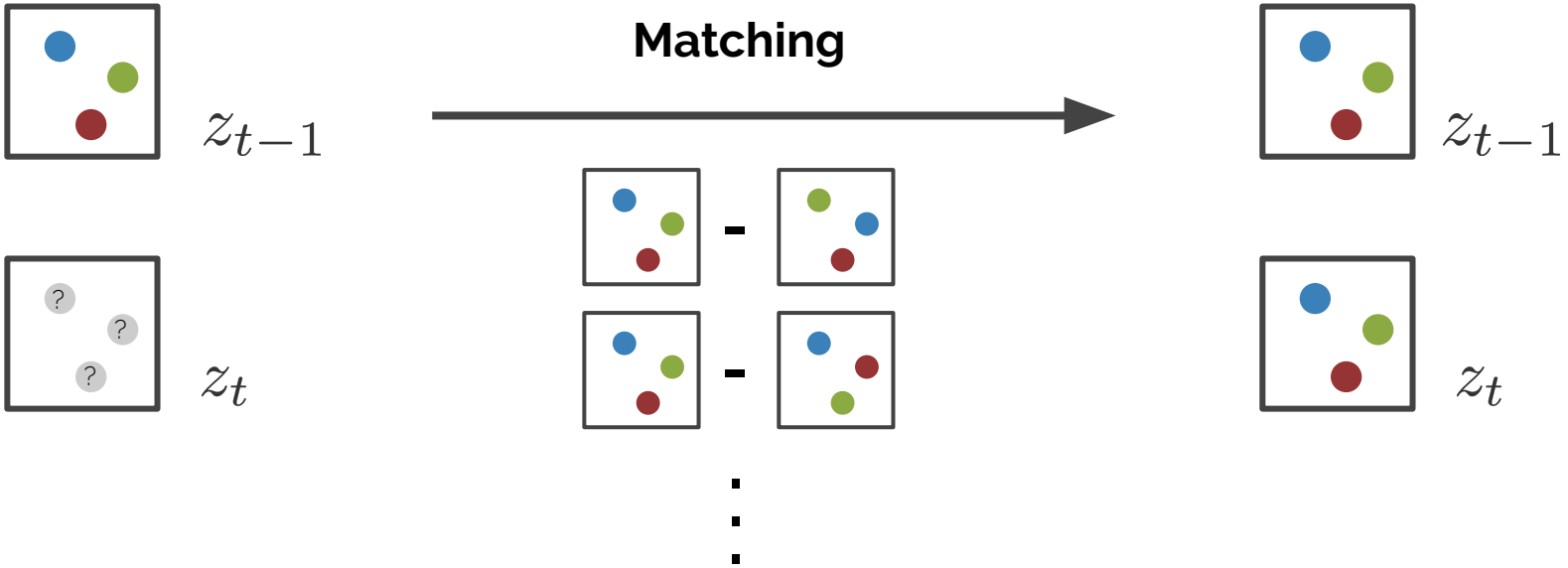
$$\log p(x_{1:T}) \geq \mathbb{E}_{q(z_{1:T}|x_{1:T})}\left[\log \frac{p(x_{1:T}, z_{1:T})}{q(z_{1:T} \mid x_{1:T})}\right] \propto \mathbb{E}_{q(z_{1:T}|x_{1:T})}\left[\sum_{t=1}^{T} \log \left\{ \frac{p(x_t \mid z_t)}{q(z_t \mid x_t)} \frac{p(z_t \mid z_{t-1})}{q(z_t \mid z_{t-1})} \right\}\right]$$



Kossen et al., Structured Object-Aware Physics Prediction for Video Modeling and Planning

The *STOVE* Model

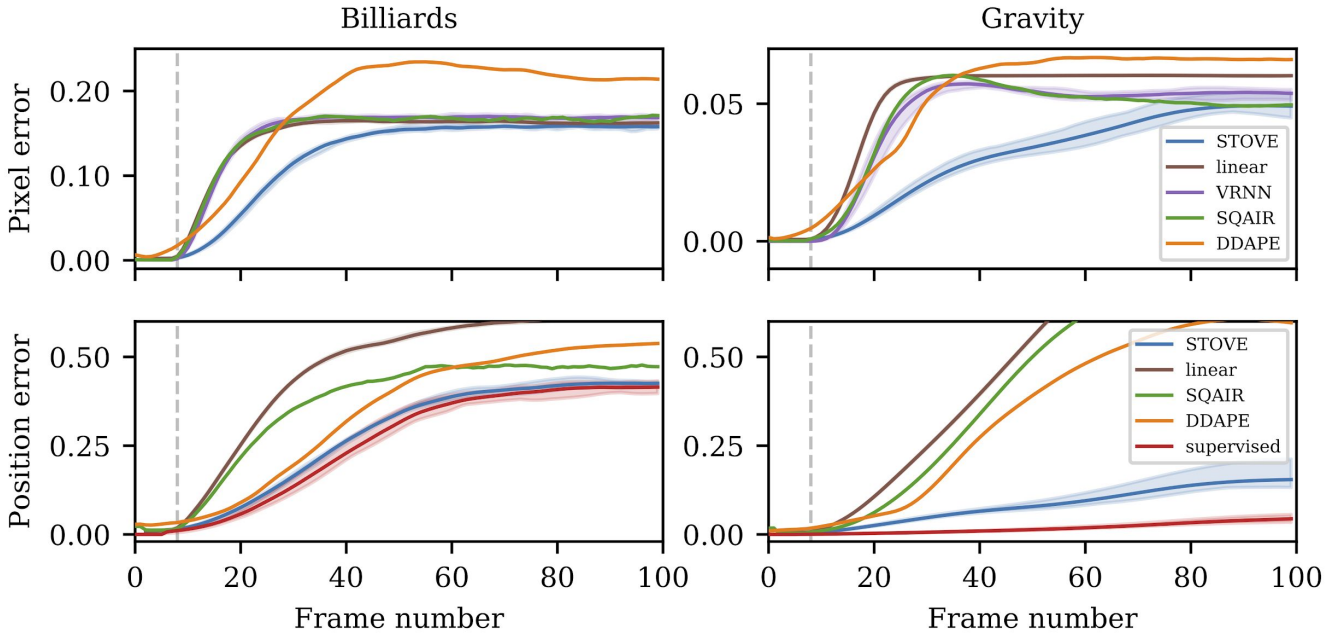
State Matching



Kossen et al., Structured Object-Aware Physics Prediction for Video Modeling and Planning

Experiments

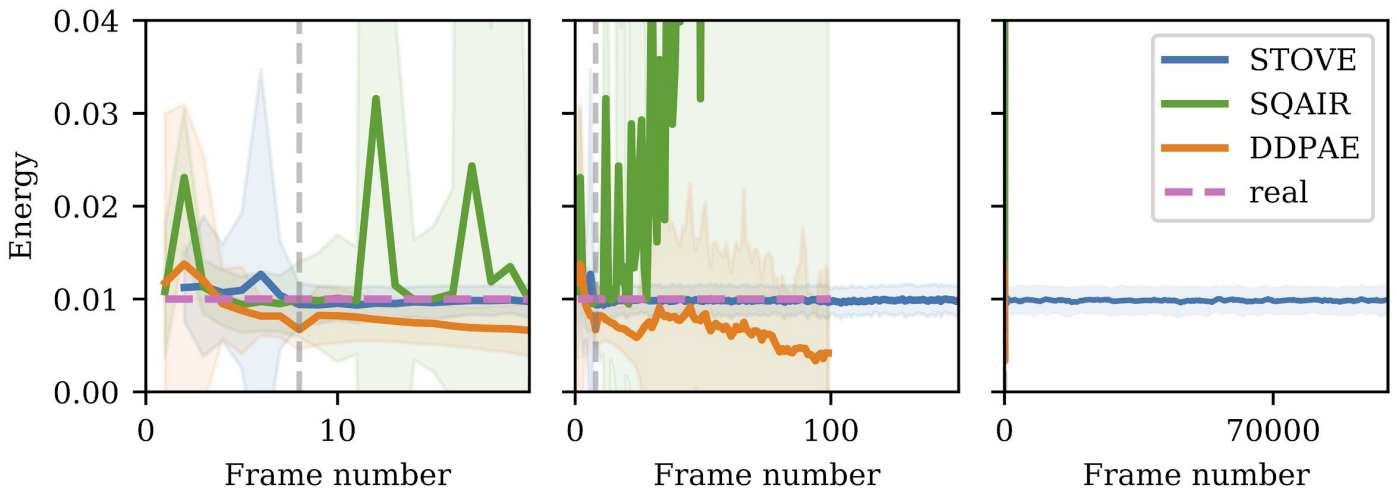
Rollout Error



Kossen et al., Structured Object-Aware Physics Prediction for Video Modeling and Planning

Experiments

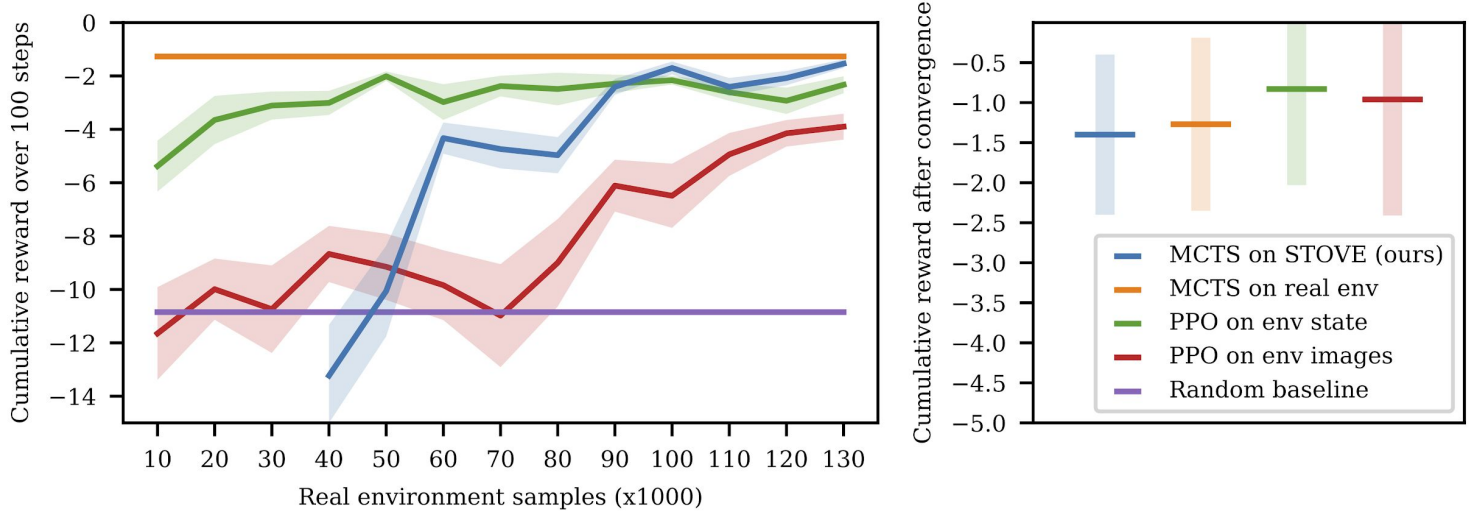
Energy Conservation



Kossen et al., Structured Object-Aware Physics Prediction for Video Modeling and Planning

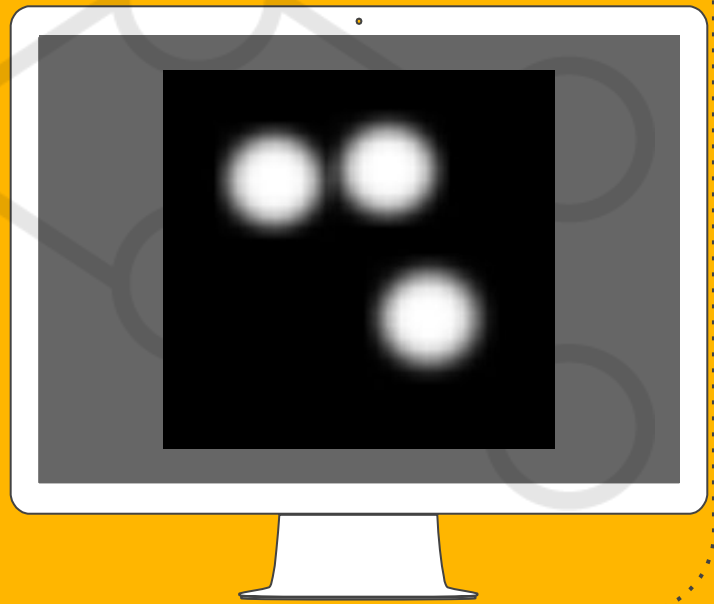
Model-Based Control

Action-Conditioned Prediction



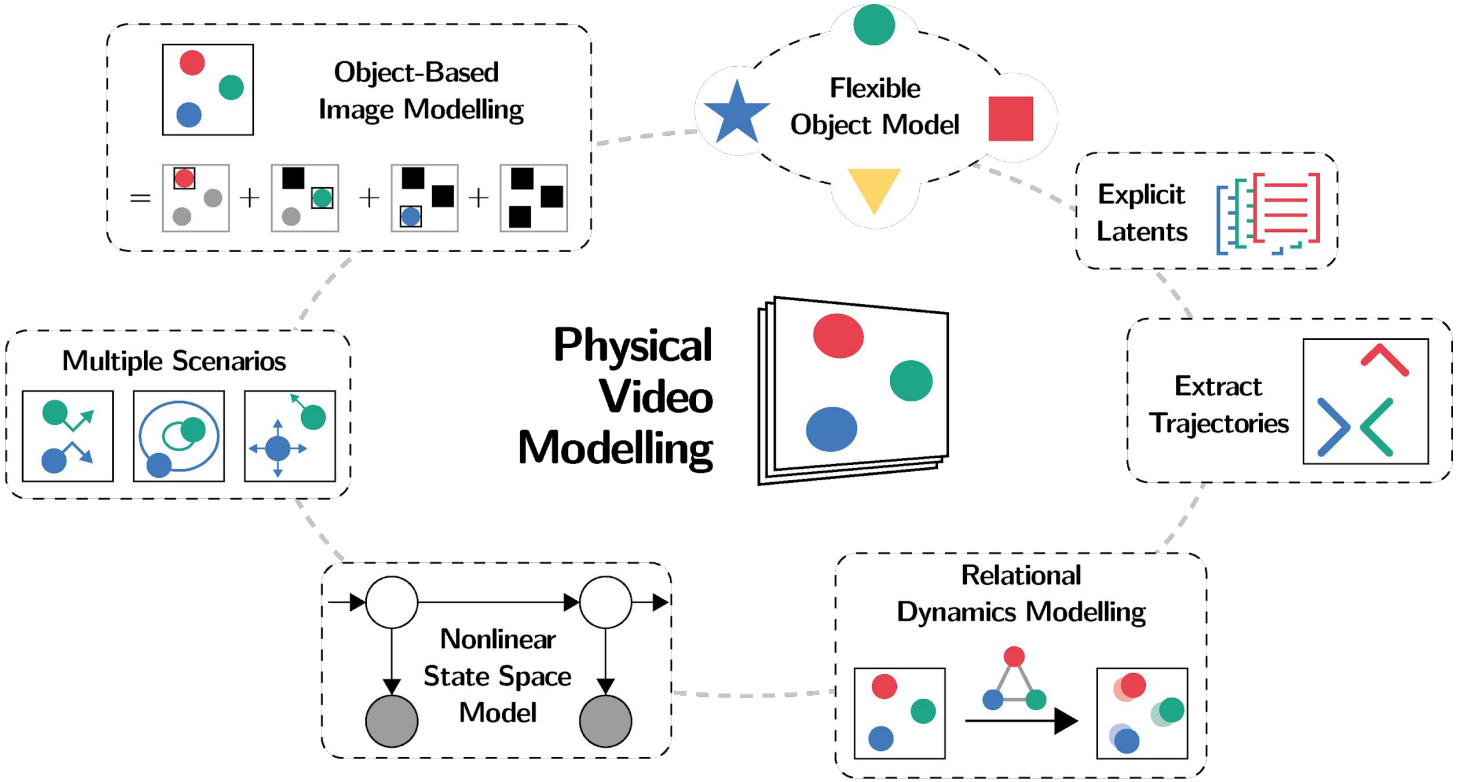
Kossen et al., Structured Object-Aware Physics Prediction for Video Modeling and Planning

Animations



Please see
github.com/jklo/stove.

STOVE – A Summary.



Kossen et al., Structured Object-Aware Physics Prediction for Video Modeling and Planning