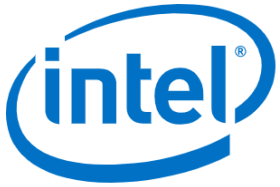
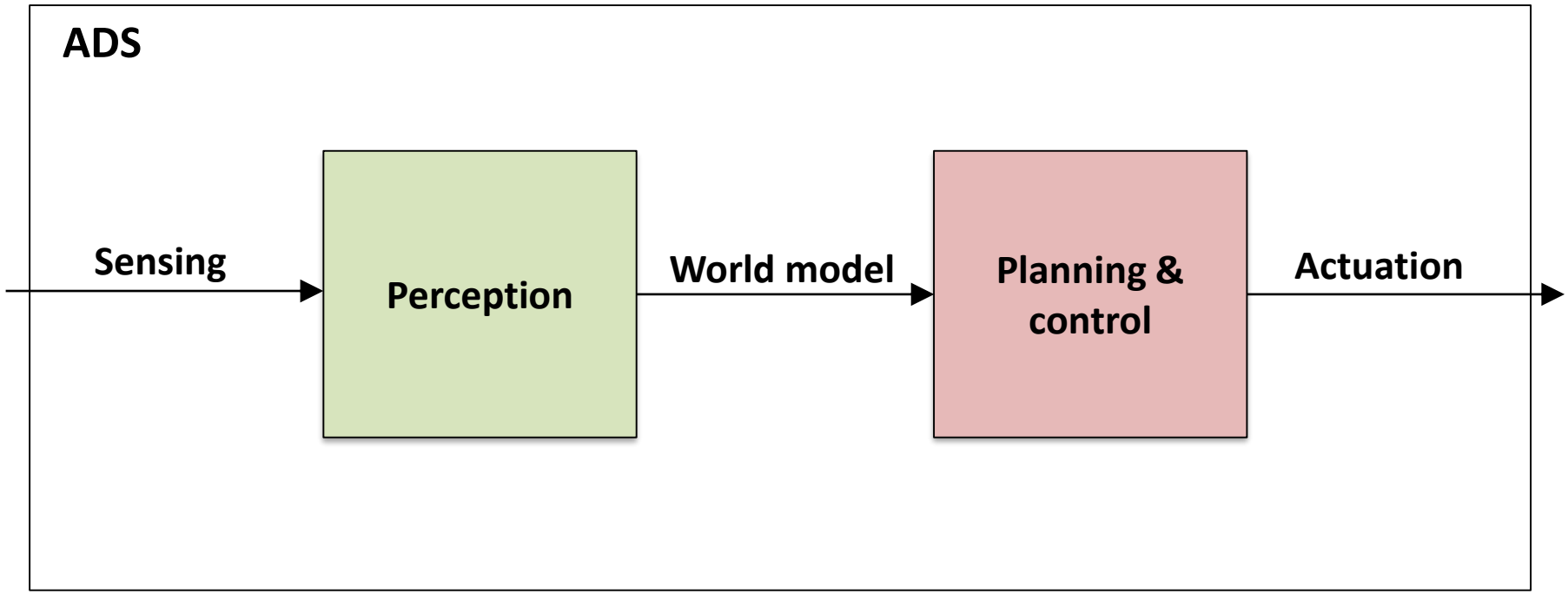
PURSS: Towards Perceptual Uncertainty Aware Responsibility Sensitive Safety with ML

Rick Salay,[1] Krzysztof Czarnecki,[1] Ignacio Alvarez,[2] Maria Soledad Elli,[2] Sean Sedwards,[1] Jack Weast[2]

[1]Dept. Electrical and Computer Engineering, Univ. of Waterloo
[2]Intel Corporation, Automated Driving Group

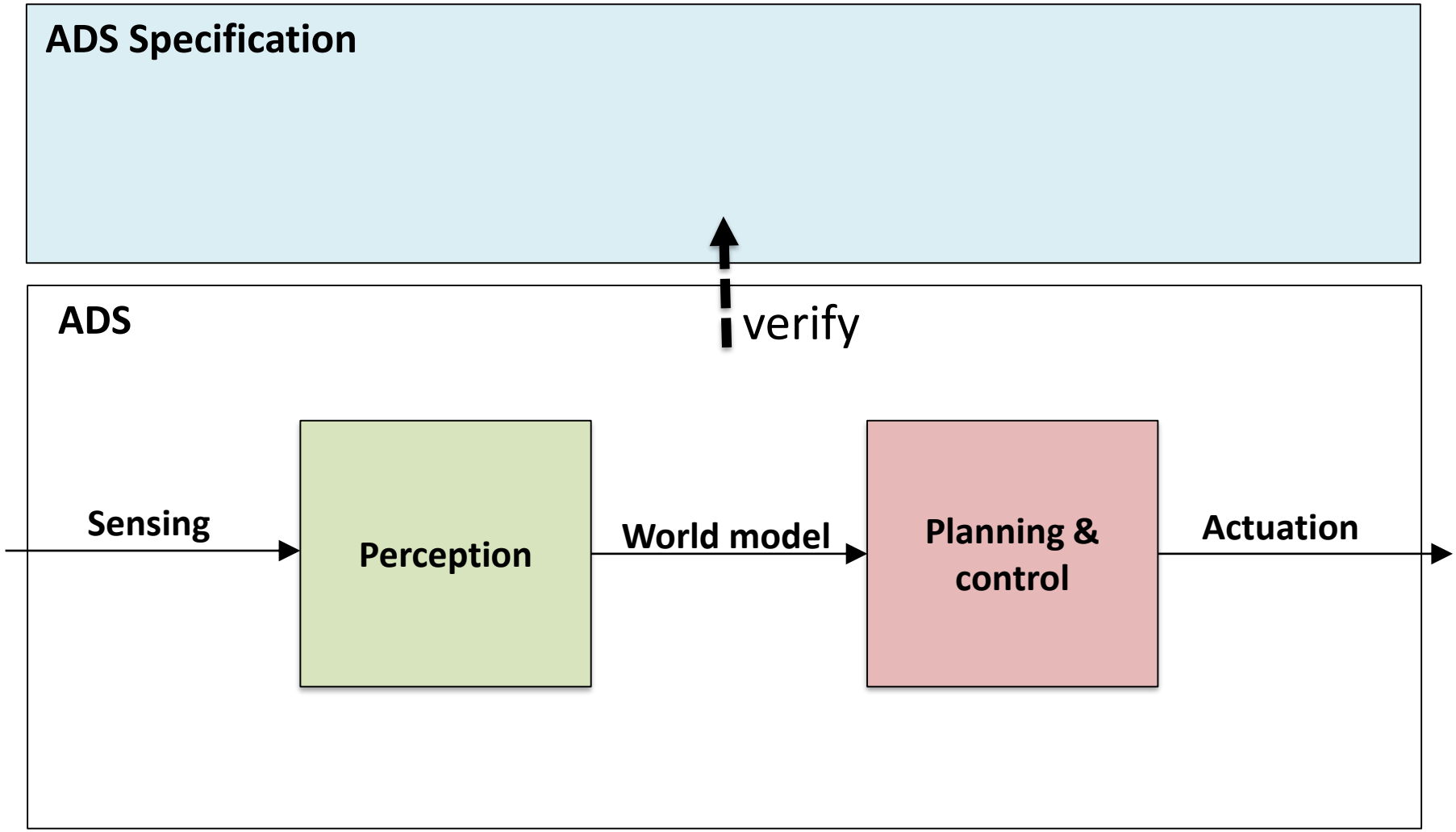


Automated Driving Systems (ADS)



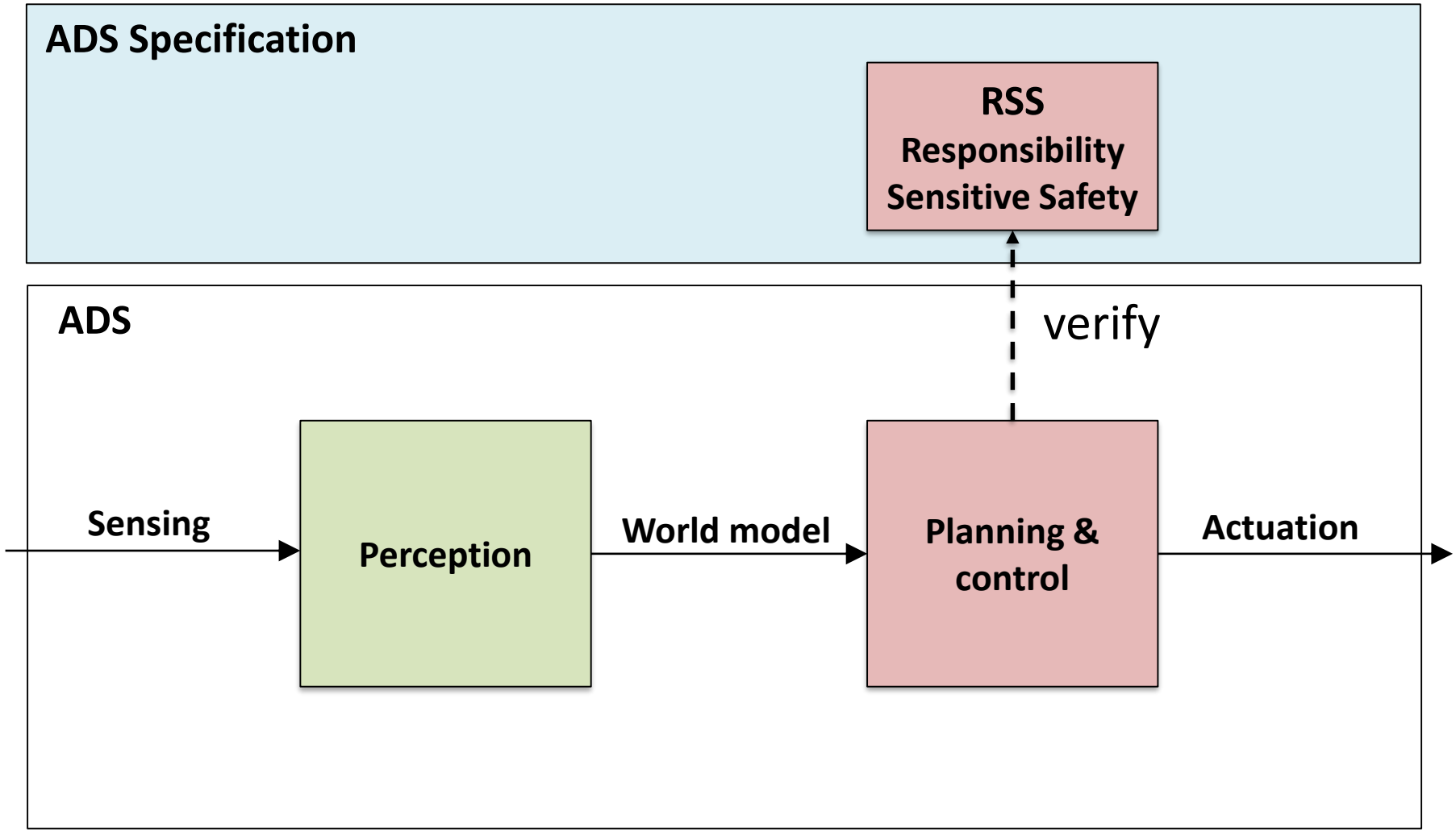
Automated Driving Systems (ADS)

Traditional Safety Assurance



Automated Driving Systems (ADS)

Traditional Safety Assurance

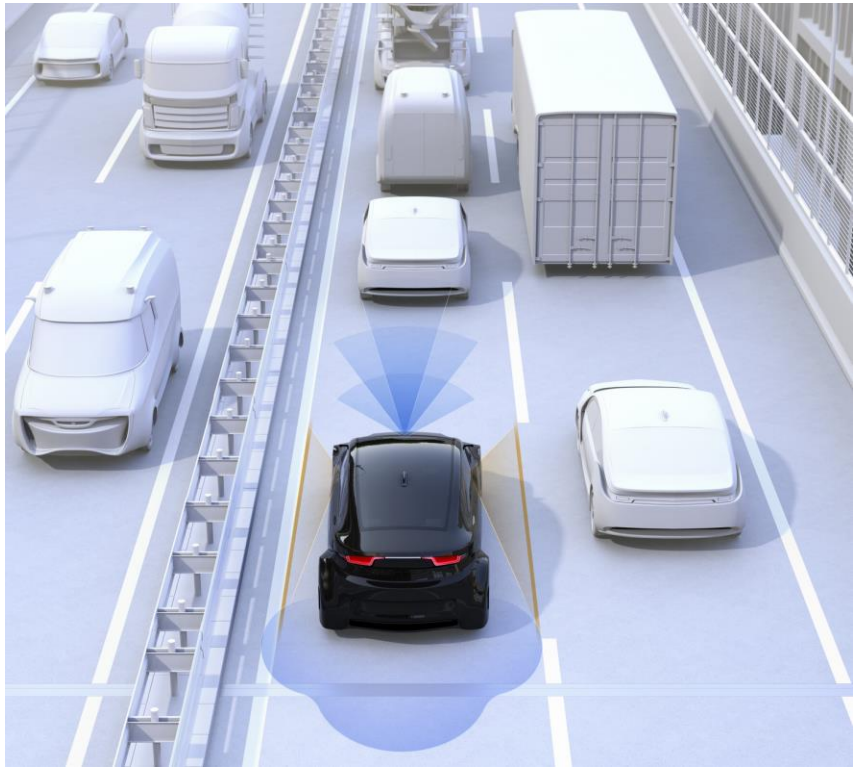


Responsible Sensitive Safety (RSS)

Formalizes

“common sense safety”

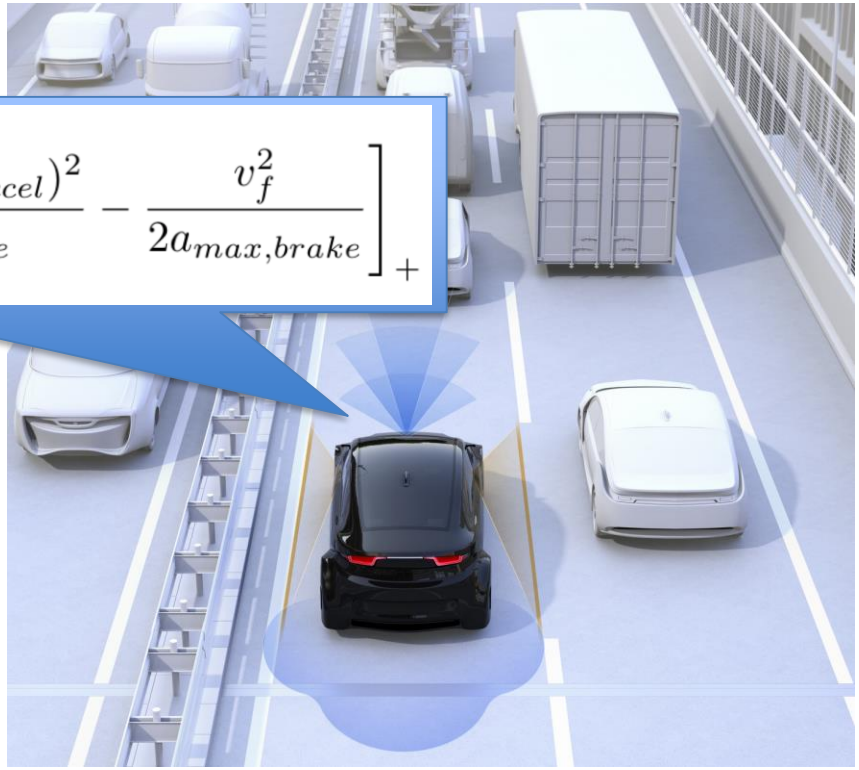
e.g., Do not hit the car in front



Responsible Sensitive Safety (RSS)

Do not hit the car in front

$$d_{min} = \left[v_r \rho + \frac{1}{2} a_{max,accel} \rho^2 + \frac{(v_r + \rho a_{max,accel})^2}{2a_{min,brake}} - \frac{v_f^2}{2a_{max,brake}} \right]_+$$



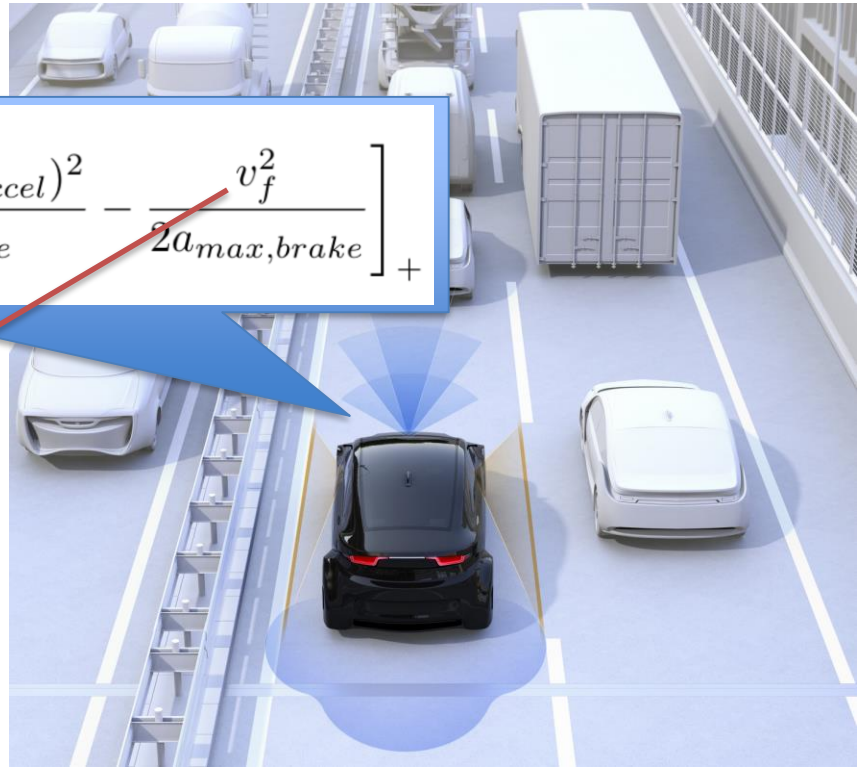
- Safe actions maintain distance d_{min}
- If d_{min} is breached, “proper response” is safe action

Shalev-Shwartz, Shai, Shaked Shammah, and Amnon Shashua. "On a formal model of safe and scalable self-driving cars." *arXiv preprint arXiv:1708.06374* (2017).

Responsible Sensitive Safety (RSS)

Do not hit the car in front

$$d_{min} = \left[v_r \rho + \frac{1}{2} a_{max,accel} \rho^2 + \frac{(v_r + \rho a_{max,accel})^2}{2a_{min,brake}} - \frac{v_f^2}{2a_{max,brake}} \right]_+$$

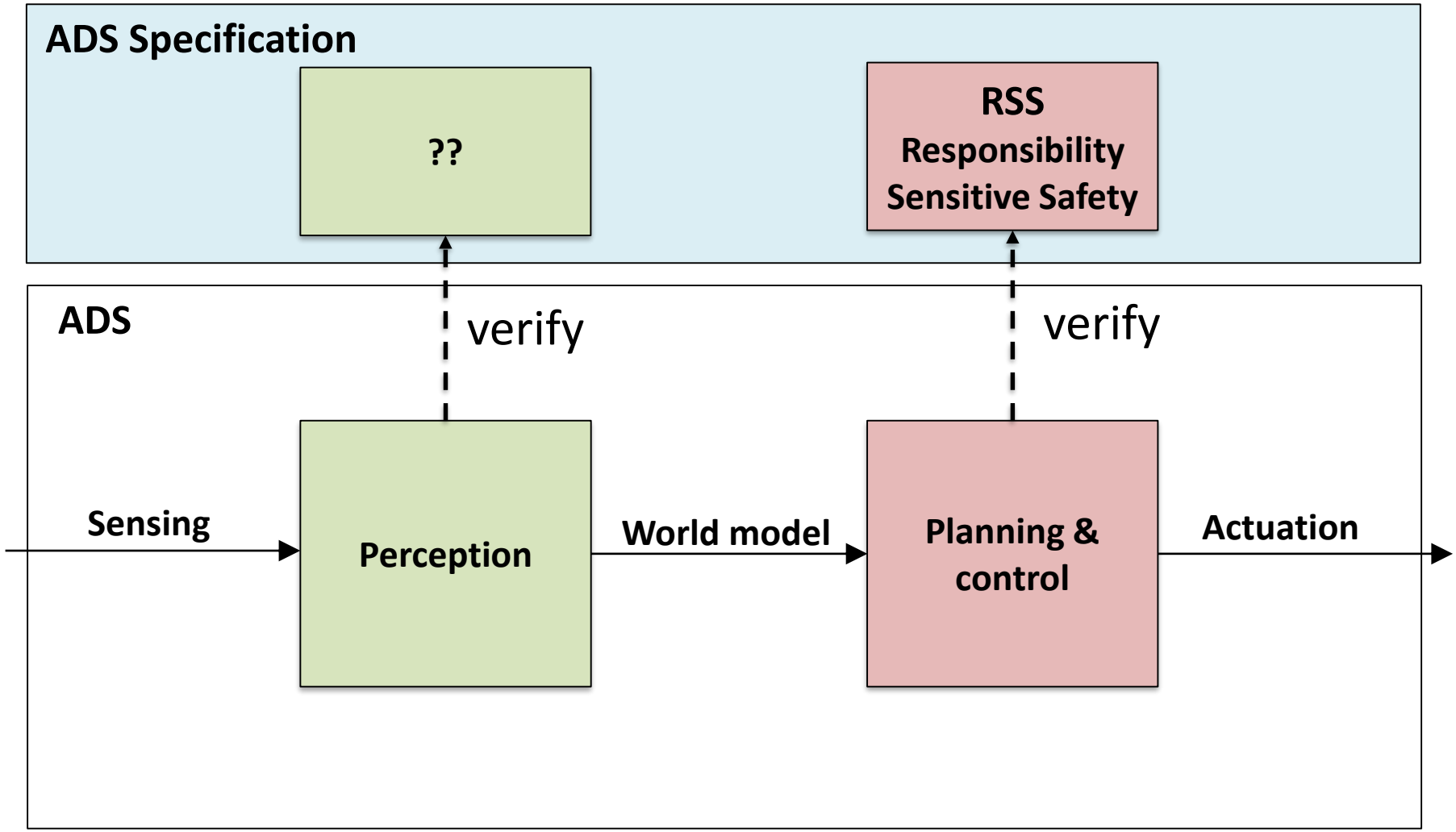


Problem: Assumes perfect perception

Misperception -> wrong action
-> safety risk!

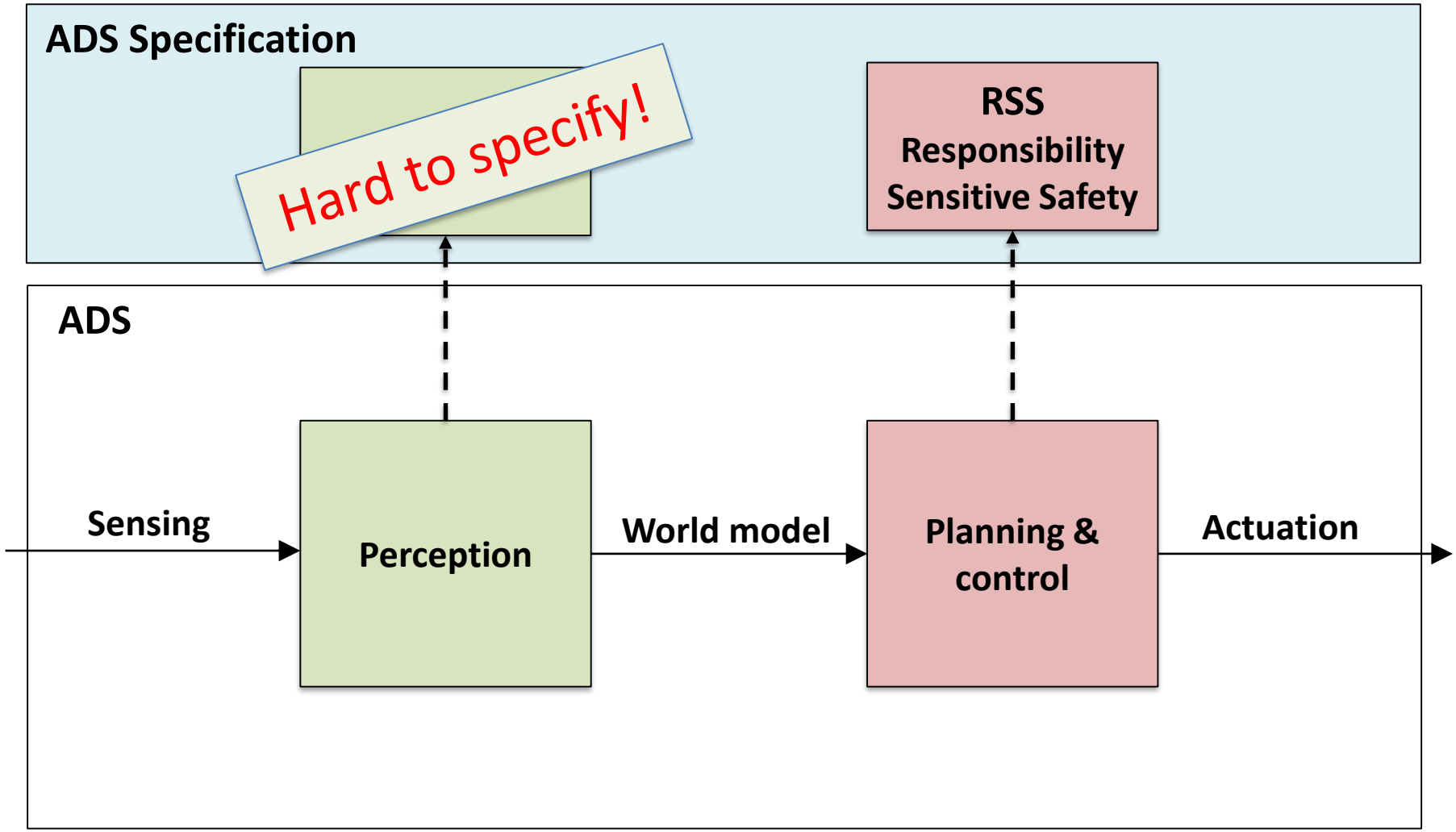
Automated Driving Systems (ADS)

Traditional Safety Assurance



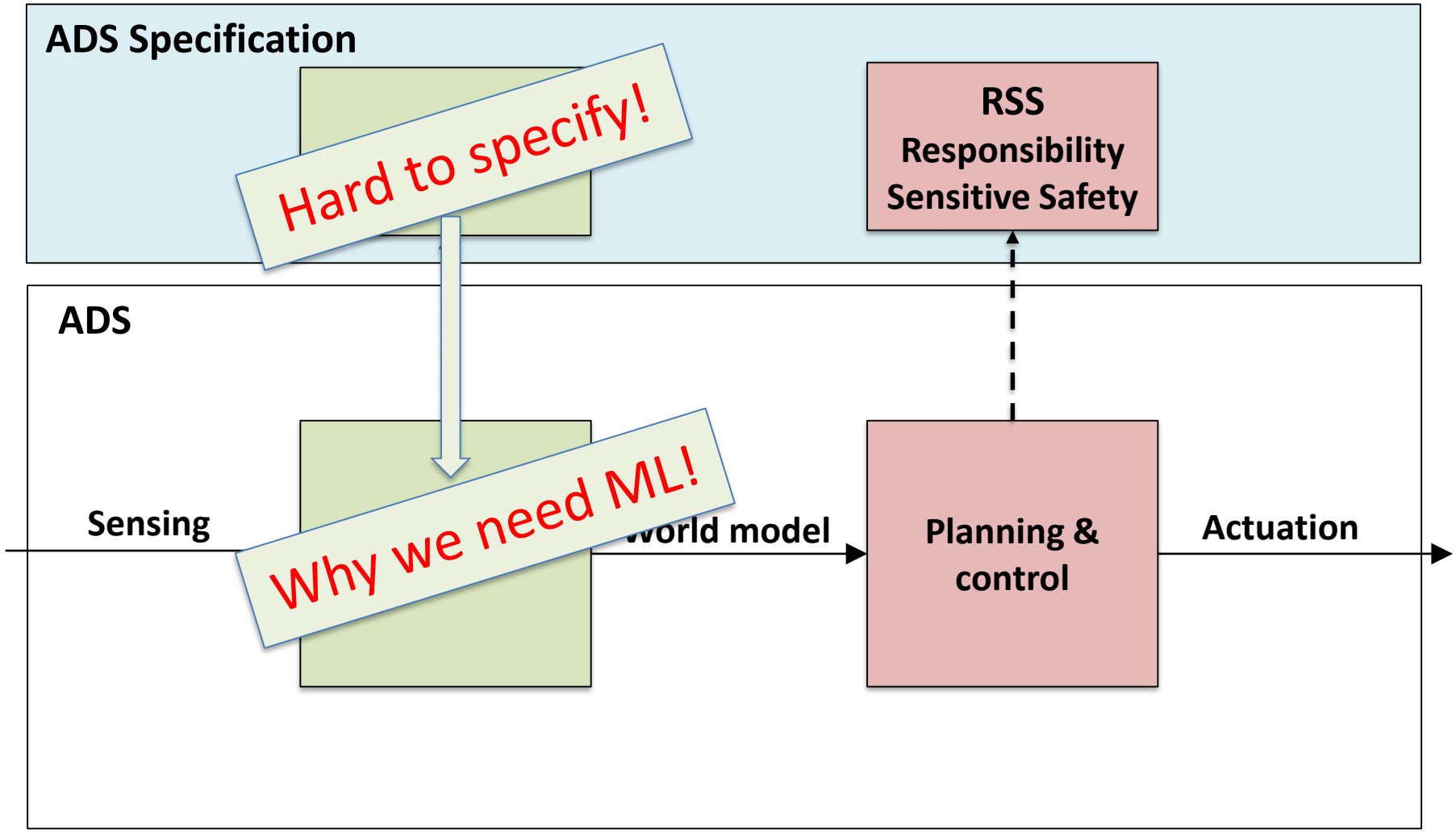
Automated Driving Systems (ADS)

Traditional Safety Assurance



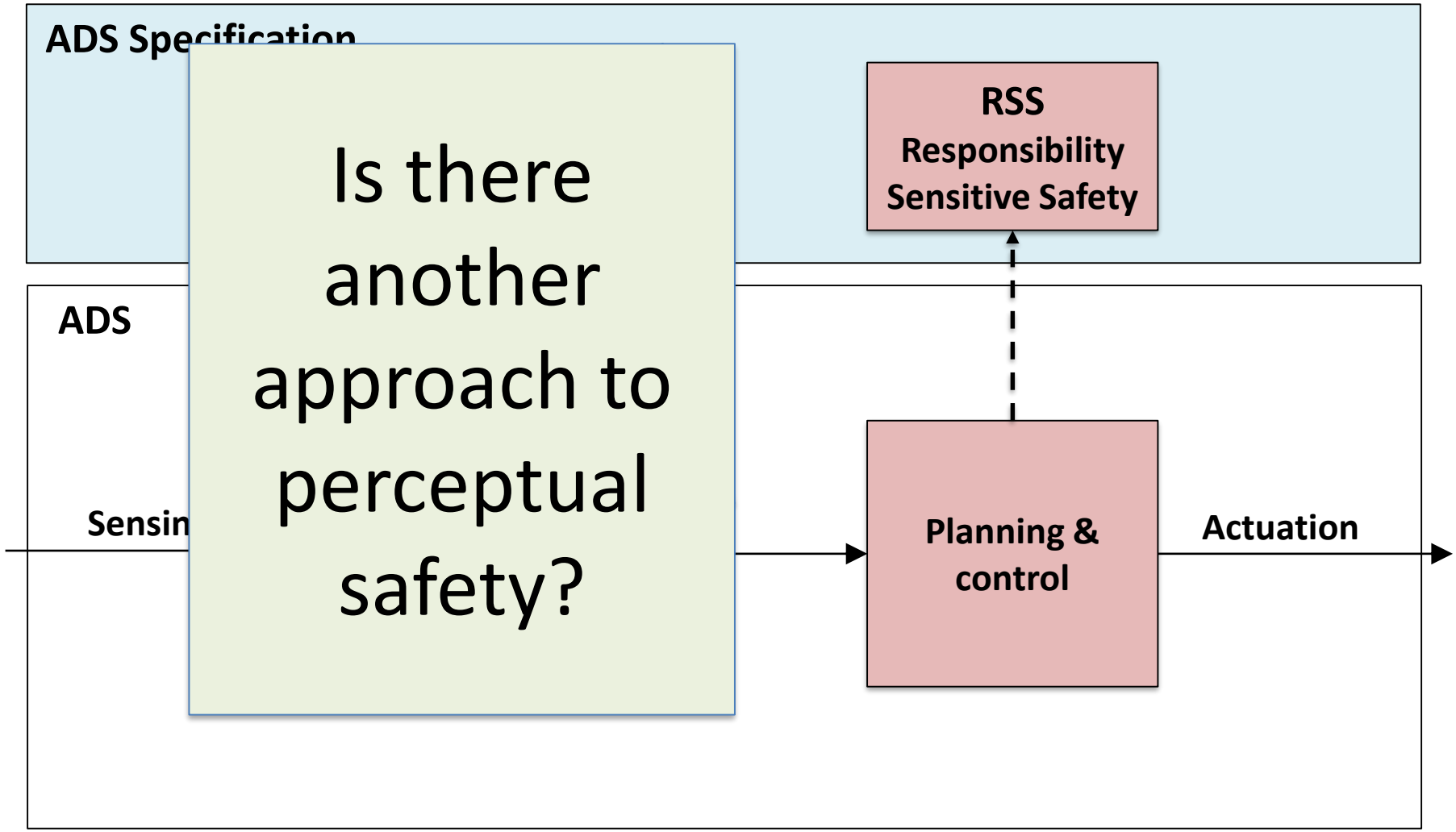
Automated Driving Systems (ADS)

Traditional Safety Assurance



Automated Driving Systems (ADS)

Traditional Safety Assurance



Is there another approach to perceptual safety?

Perceptual Uncertainty

- Uncertainty of perceptual component is cause of misperception
  - many factors*: poor labeling, inadequate dataset coverage, etc.
- ML components can report their own uncertainty!
  - as long as they are calibrated…

*Czarnecki, Krzysztof, and Rick Salay. "Towards a framework to manage perceptual uncertainty for safe automated driving." In *International Conference on Computer Safety, Reliability, and Security*, pp. 439-445. Springer, Cham, 2018.

PURSS

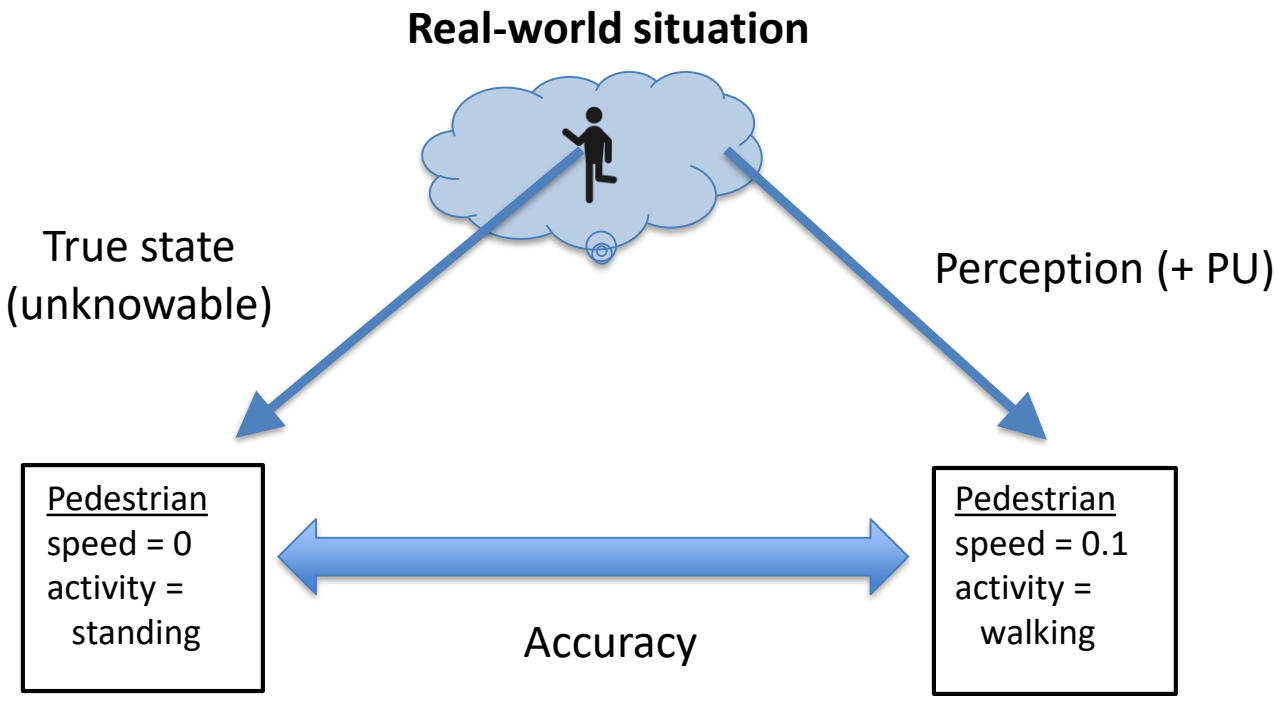
PURSS = perceptual uncertainty (PU) + RSS

Safety Idea:

Use perceptual uncertainty measure to make RSS rules appropriately cautious and limit safety risk

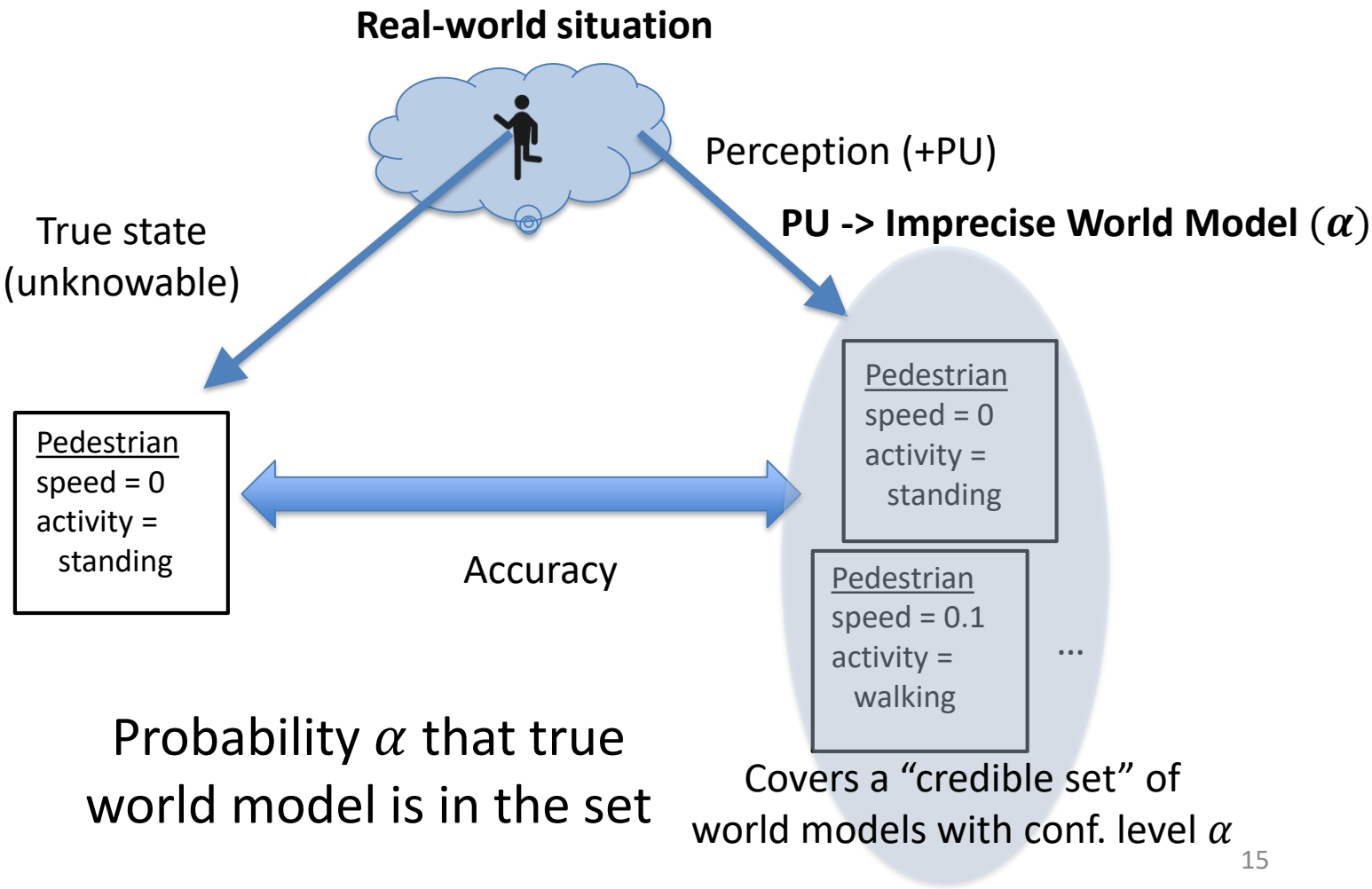
PURSS formalizes this idea

Precise World Model

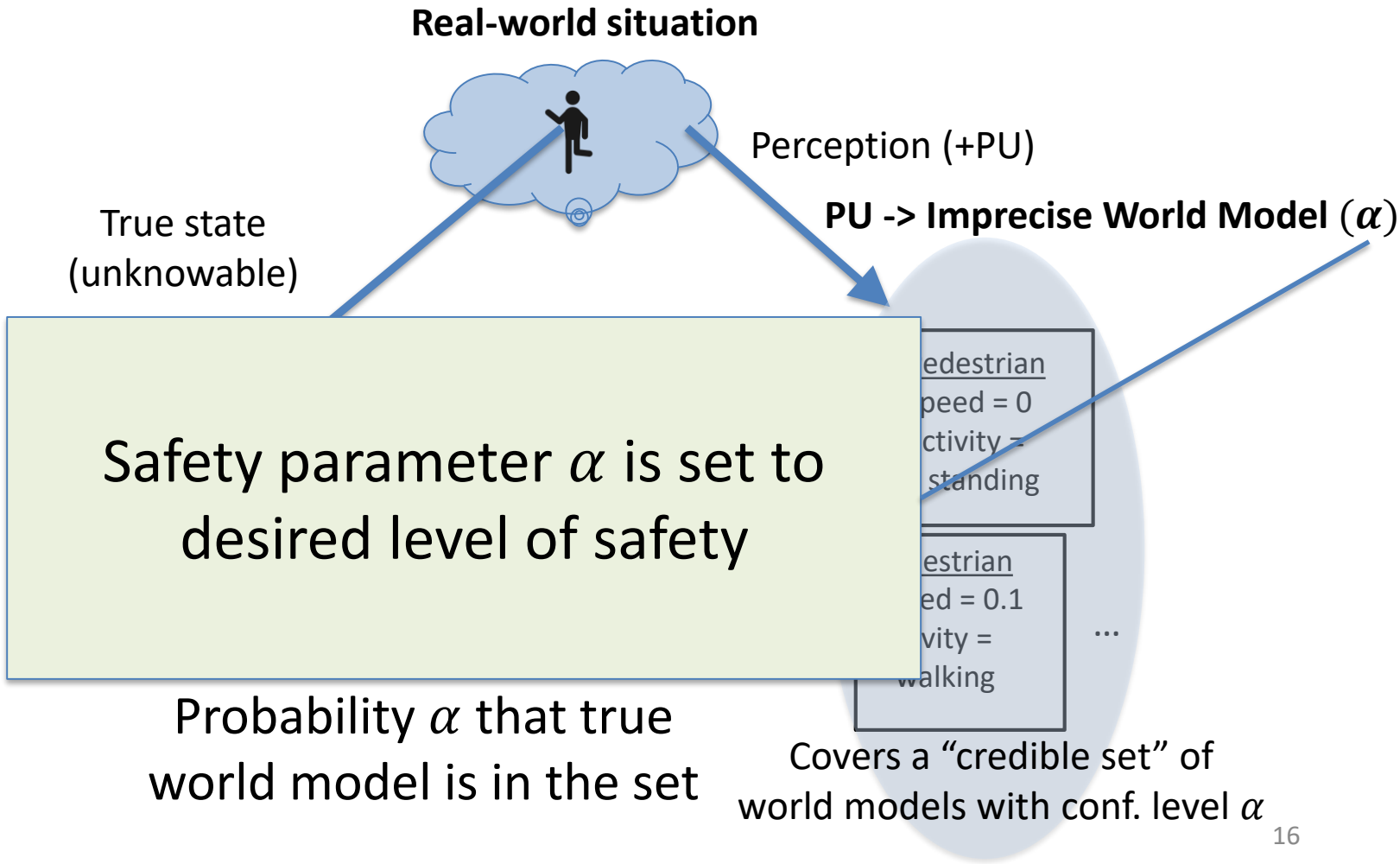


Misperception: precise but inaccurate

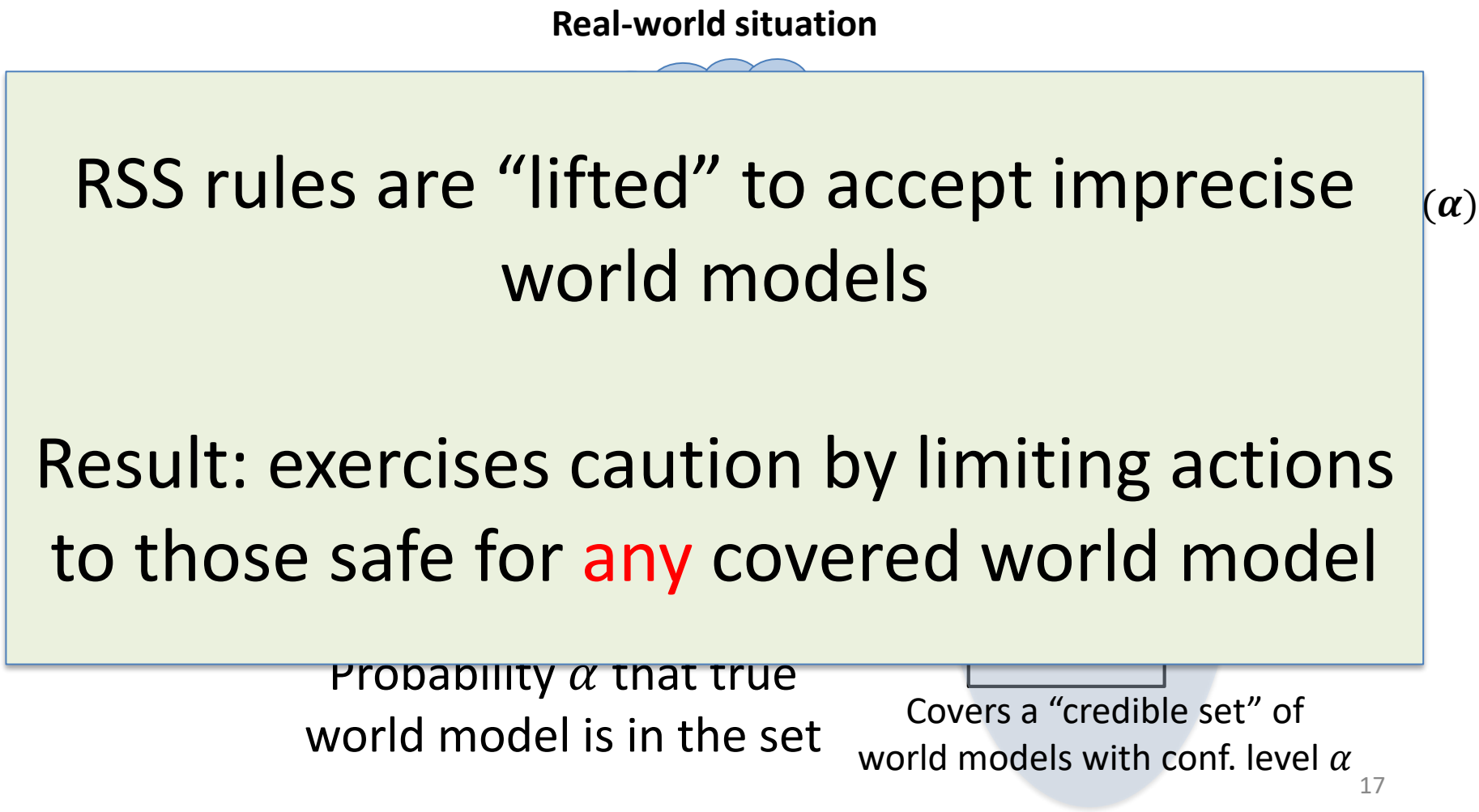
Perceptual Uncertainty Handling via Imprecise World Models



Perceptual Uncertainty Handling via Imprecise World Models



Perceptual Uncertainty Handling via Imprecise World Models

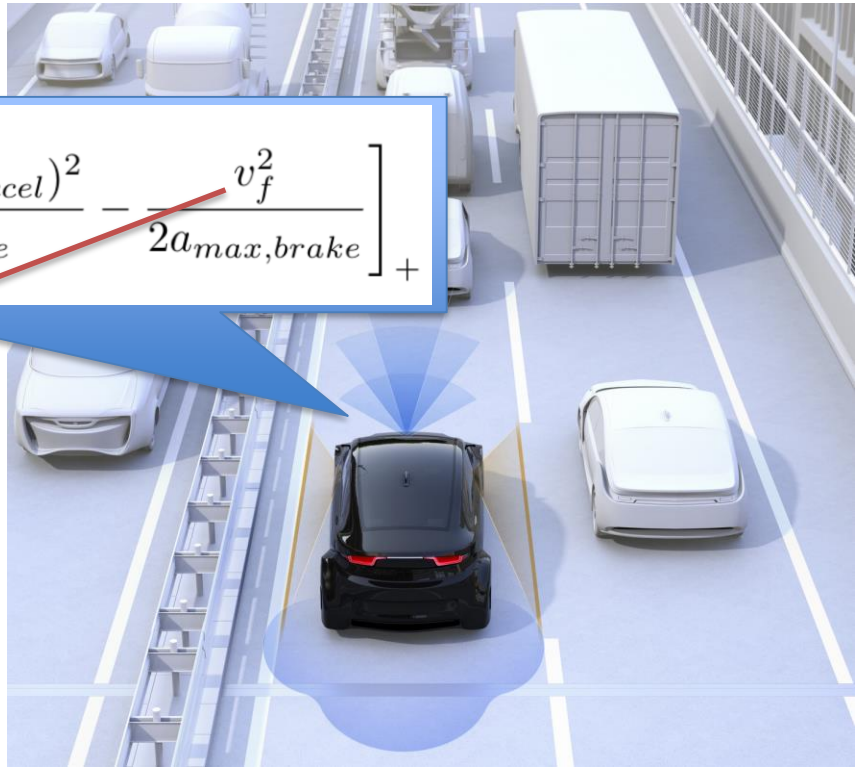


Responsible Sensitive Safety (RSS)

Do not hit the car in front

$$d_{min} = \left[v_r \rho + \frac{1}{2} a_{max,accel} \rho^2 + \frac{(v_r + \rho a_{max,accel})^2}{2a_{min,brake}} - \frac{v_f^2}{2a_{max,brake}} \right]_+$$

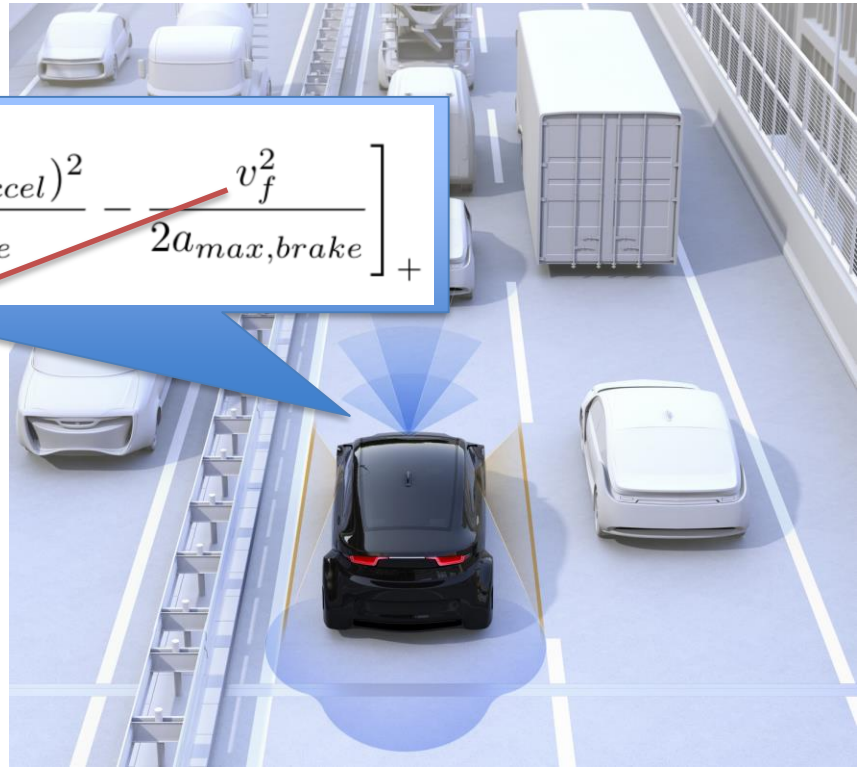
Lifting: replace values with credible intervals corresponding to α



Responsible Sensitive Safety (RSS)

Do not hit the car in front

$$d_{min} = \left[v_r \rho + \frac{1}{2} a_{max,accel} \rho^2 + \frac{(v_r + \rho a_{max,accel})^2}{2a_{min,brake}} - \frac{v_f^2}{2a_{max,brake}} \right]_+$$



Lifting: replace values with credible intervals corresponding to α

e.g.,

precise: $v_f = 30\ m/s$

PU: $\sigma_f^2 = 1\ m/s$

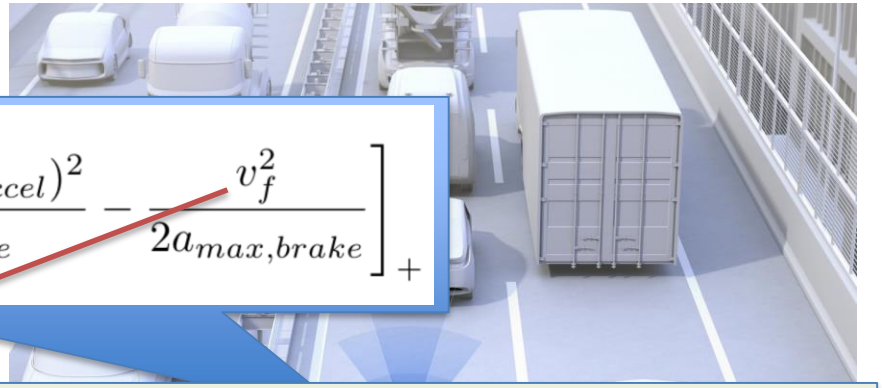
lift to imprecise:

$\alpha = 68\%$: $v_f = [29,31]\ m/s$

$\alpha = 95\%$: $v_f = [28,32]\ m/s$

Responsible Sensitive Safety (RSS)

Do not hit the car in front



$$d_{min} = \left[v_r \rho + \frac{1}{2} a_{max,accel} \rho^2 + \frac{(v_r + \rho a_{max,accel})^2}{2a_{min,brake}} - \frac{v_f^2}{2a_{max,brake}} \right]_+$$

Lifting: replace values with cre
intervals corresponding to

e.g.,

precise: $v_f = 30\ m/s$

PU: $\sigma_f^2 = 1\ m/s$

lift to imprecise:

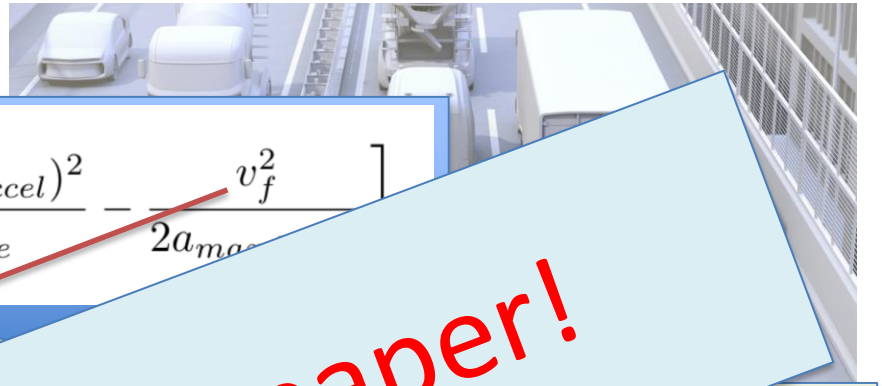
$\alpha = 68\%$: $v_f = [29,31]\ m/s$

$\alpha = 95\%$: $v_f = [28,32]\ m/s$

Given uncertainty σ_f^2, increasing confidence $\alpha \Rightarrow$ decreasing precision of $v_f \Rightarrow$ larger d_{min} to be more cautious

Responsible Sensitive Safety (RSS)

Do not hit the car in front



$$d_{min} = \left[v_r \rho + \frac{1}{2} a_{max,accel} \rho^2 + \frac{(v_r + \rho a_{max,accel})^2}{2a_{min,brake}} - \frac{v_f^2}{2a_{ma}} \right]$$

Lifting: replace values wit
intervals corr

e.g.,

li

$v_f = [29,31]\ m/s$

95%: $v_f = [28,32]\ m/s$

v_f^2,

confidence $\alpha \Rightarrow$

easing precision of $v_f \Rightarrow$

larger d_{min} to be more cautious

See details in the paper!

Benefits and Costs

- Benefit: Safety parameter α can be increased to get as safe as you want
  - RSS rules become correspondingly more cautious
- Cost: More cautious behaviour may negatively impact progress
- Important future work: negotiating the trade-off

Summary

- RSS provides a spec on planning & control
  - supports traditional safety assurance
- Perception is hard to specify and needs ML
  - different safety approach is needed
- PURSS approach to safety
  - Set desired level of safety (α)
  - Perceptual uncertainty $\rightarrow_{\alpha}$ imprecise world models
  - Lift RSS rules to be correspondingly cautious
- Much further work coming!