The Battle for Filter Supremacy:

A Comparative Study of the Multi-State Constraint Kalman Filter and the Sliding Window Filter

Lee Clement, Valentin Peretroukhin, Jacob Lambert, and Jonathan Kelly

CRV 2015, Halifax, Canada

Institute for Aerospace Studies
UNIVERSITY OF TORONTO

STARS
LABORATORY

Motivation: Monocular Camera + IMU

Aerial vehicles

Wearable sensors

Smartphones

Motivation: Monocular Camera + IMU

Aerial vehicles

Wearable sensors

Smartphones

How can we use these sensors to **navigate** an unknown environment?

What is the **best algorithm** to use **online** in this context?

Institute for Aerospace Studies
UNIVERSITY OF TORONTO

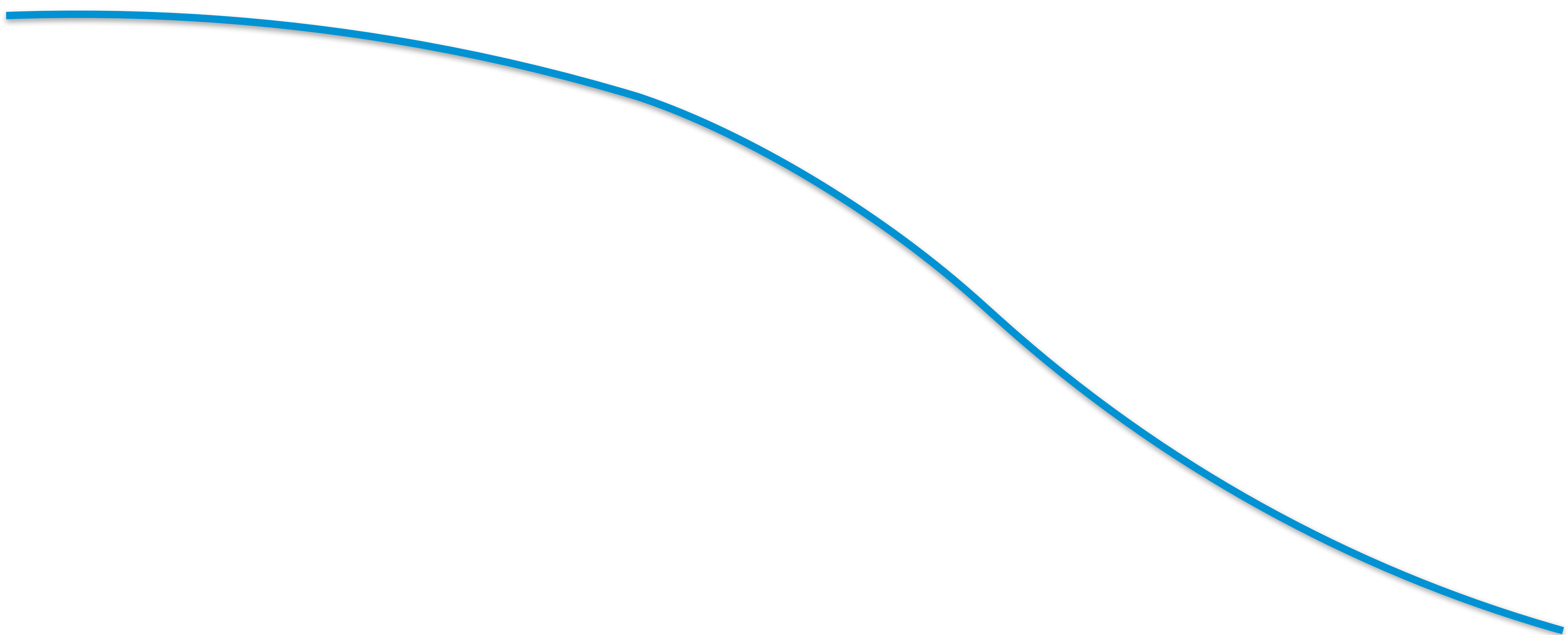
Lee Clement, Valentin Peretroukhin, Jacob Lambert, and Jonathan Kelly

Traditional Solution: Extended Kalman Filter

Goal: Use an IMU and a monocular camera to estimate motion *without a map*.

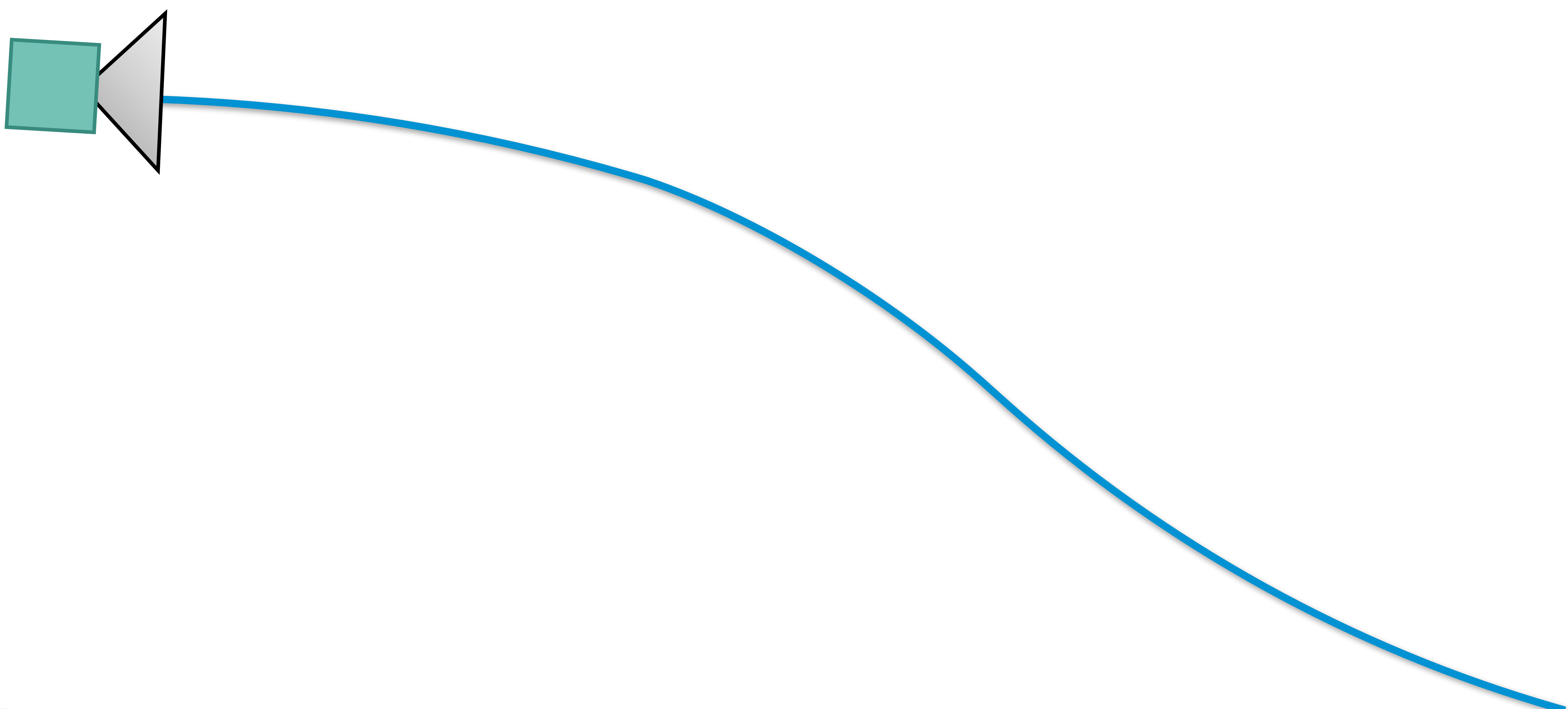
Traditional Solution: Extended Kalman Filter

Goal: Use an IMU and a monocular camera to estimate motion *without a map*.



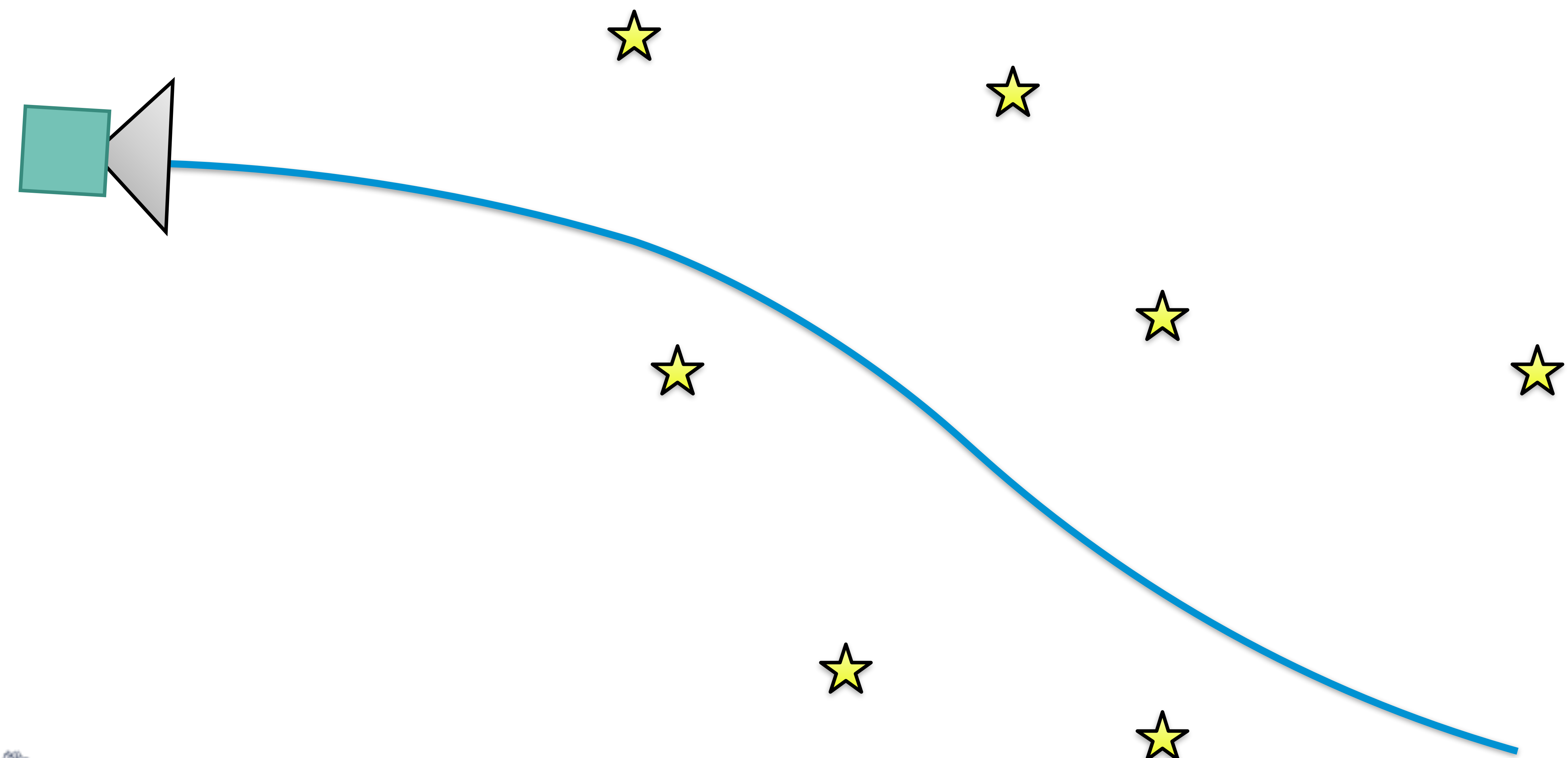
Traditional Solution: Extended Kalman Filter

Goal: Use an IMU and a monocular camera to estimate motion *without a map*.



Traditional Solution: Extended Kalman Filter

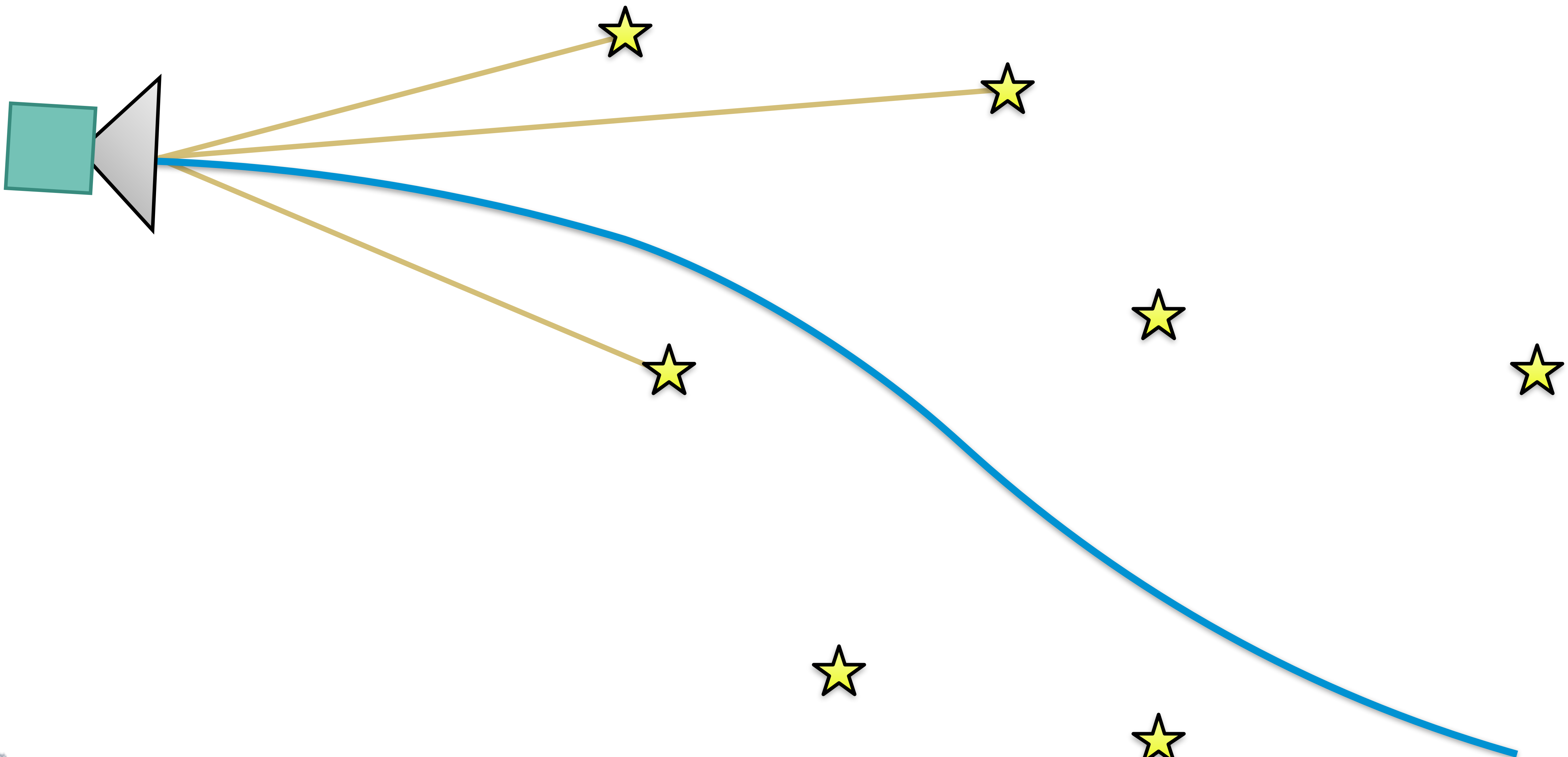
Goal: Use an IMU and a monocular camera to estimate motion *without a map*.



Institute for Aerospace Studies
UNIVERSITY OF TORONTO

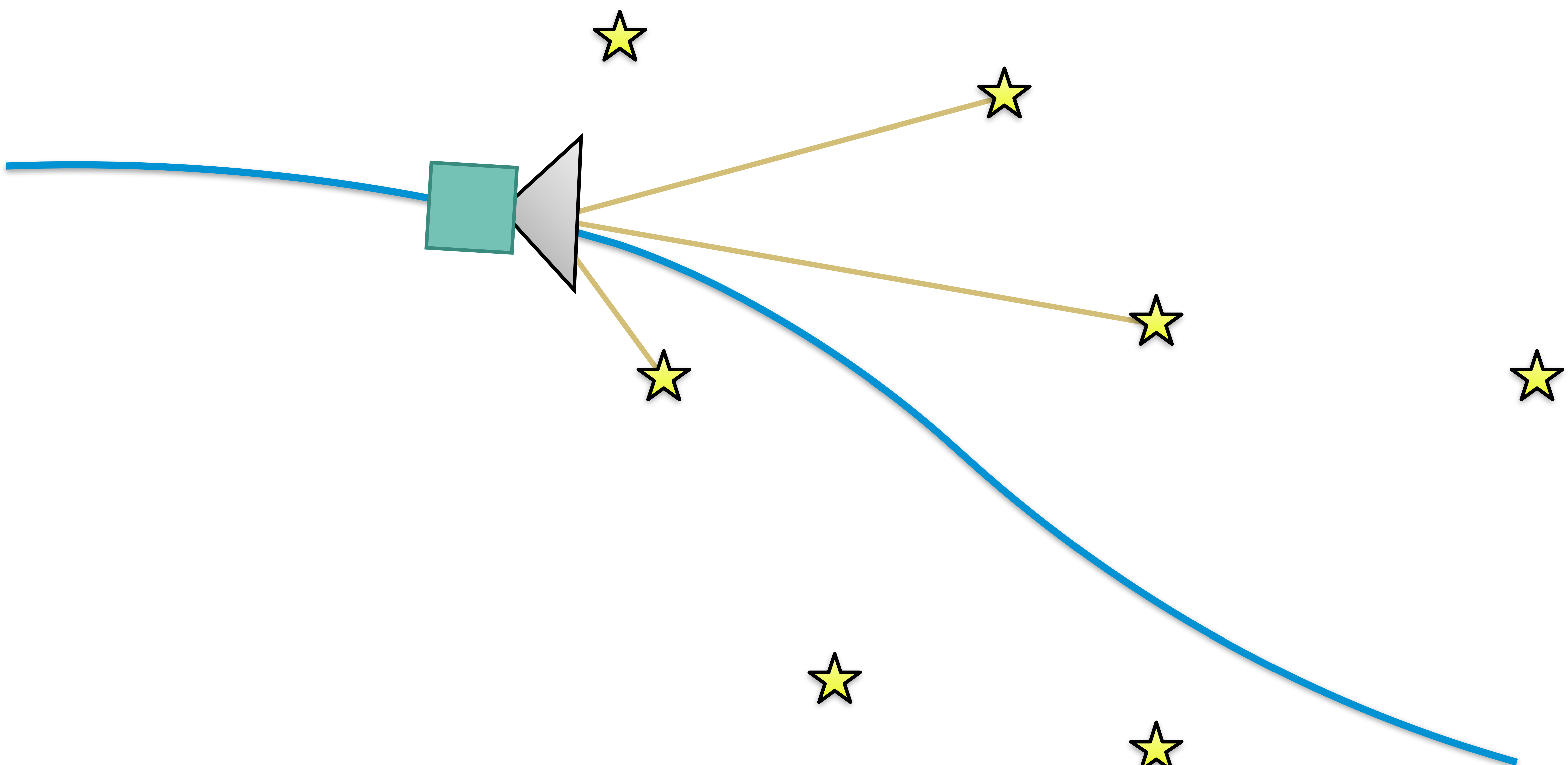
Traditional Solution: Extended Kalman Filter

Goal: Use an IMU and a monocular camera to estimate motion *without a map*.



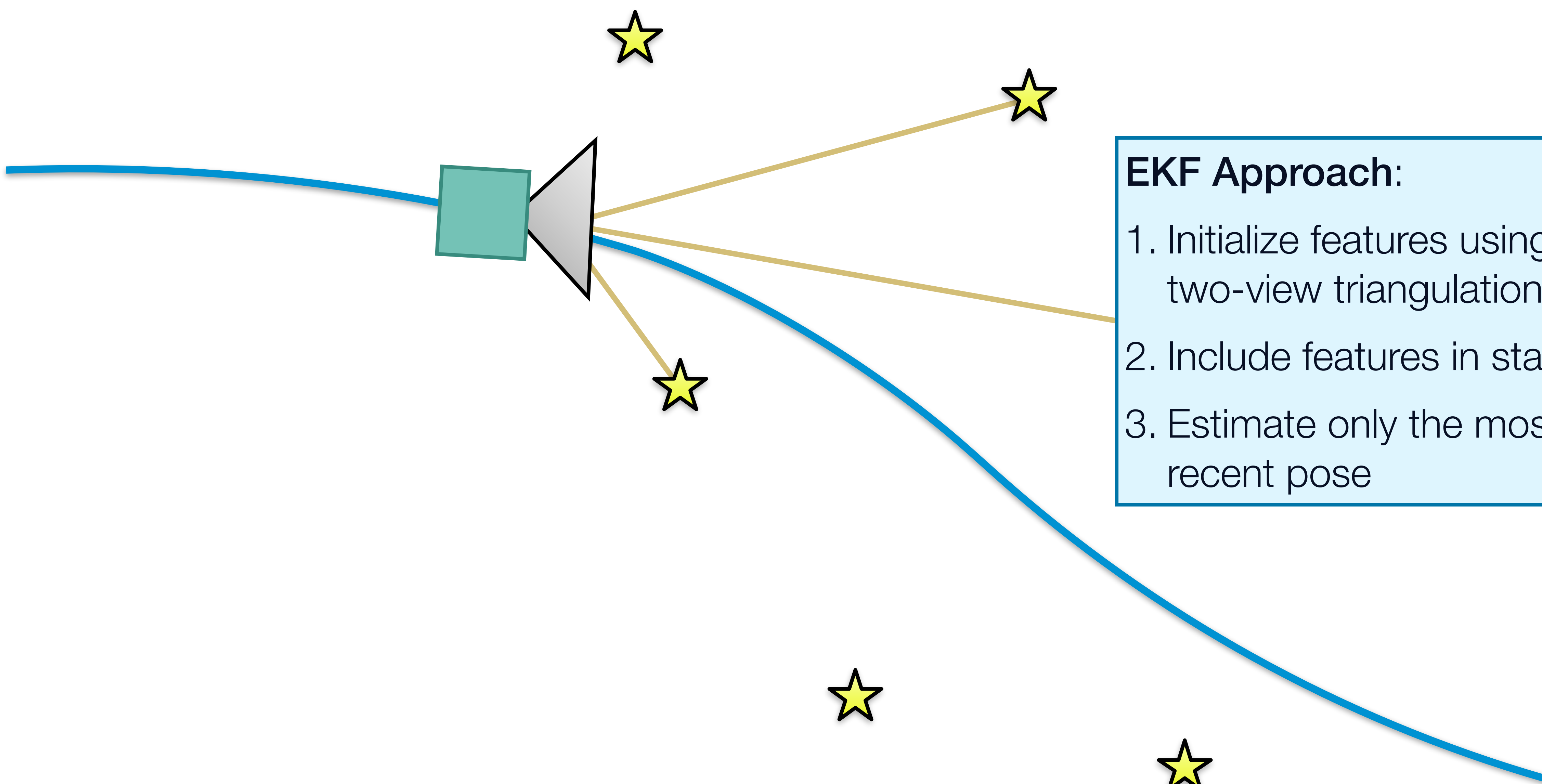
Traditional Solution: Extended Kalman Filter

Goal: Use an IMU and a monocular camera to estimate motion *without a map*.



Traditional Solution: Extended Kalman Filter

Goal: Use an IMU and a monocular camera to estimate motion *without a map*.



EKF Approach:

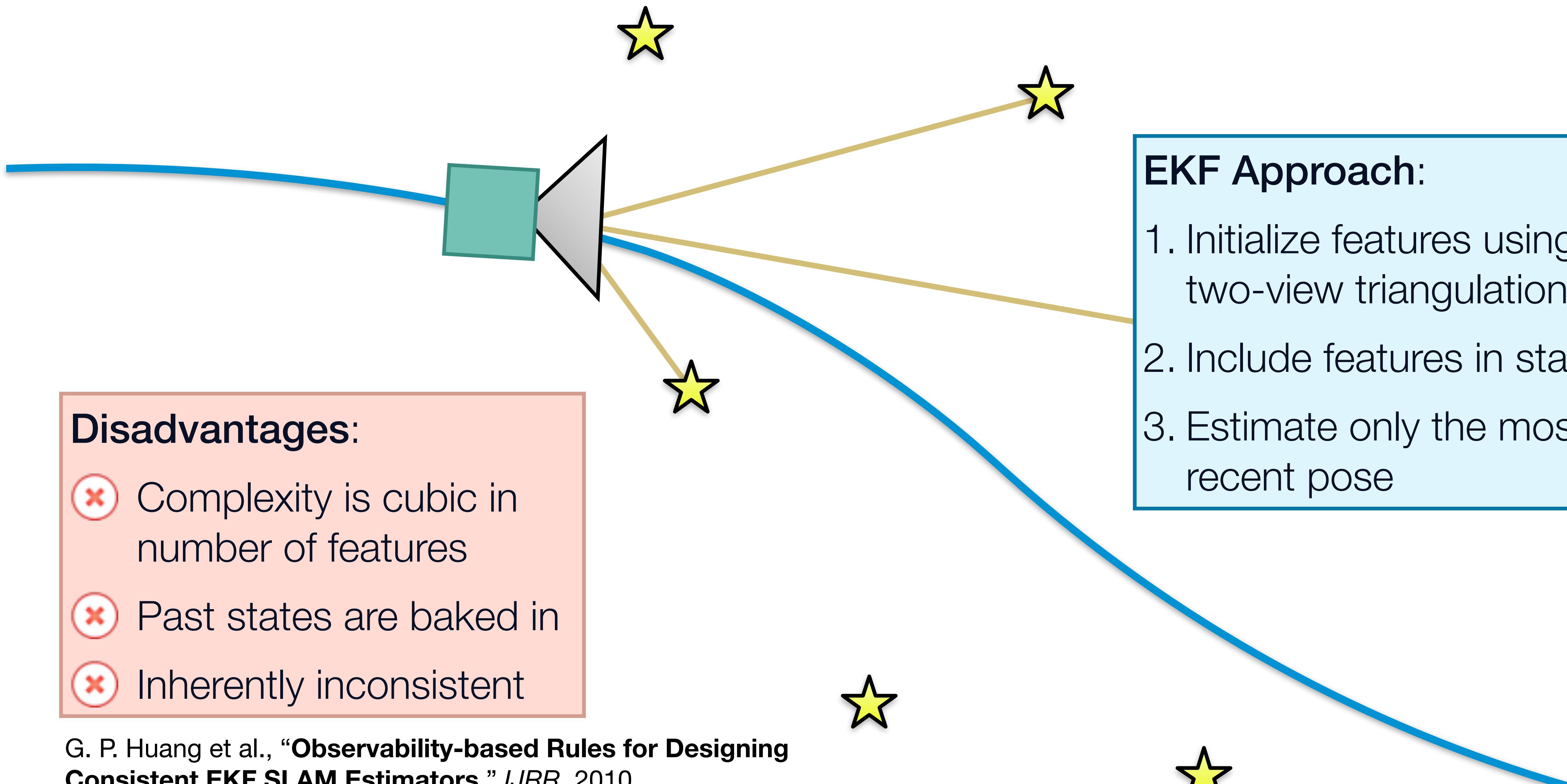
1. Initialize features using two-view triangulation
2. Include features in state
3. Estimate only the most recent pose

Institute for Aerospace Studies
UNIVERSITY OF TORONTO

Lee Clement, Valentin Peretroukhin, Jacob Lambert, and Jonathan Kelly

Traditional Solution: Extended Kalman Filter

Goal: Use an IMU and a monocular camera to estimate motion *without a map*.



EKF Approach:

1. Initialize features using two-view triangulation
2. Include features in state
3. Estimate only the most recent pose

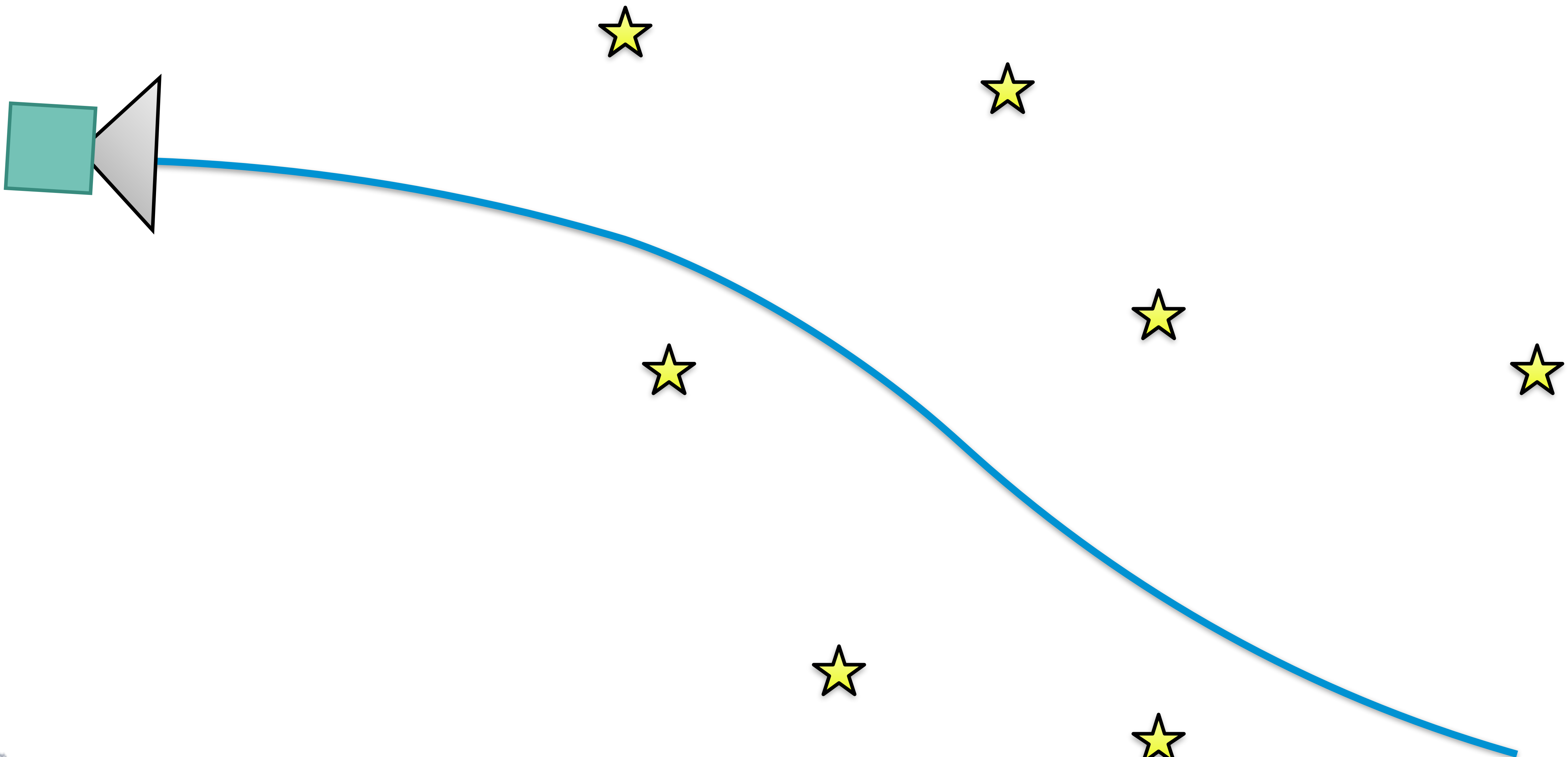
Disadvantages:

- Complexity is cubic in number of features
- Past states are baked in
- Inherently inconsistent

G. P. Huang et al., "**Observability-based Rules for Designing Consistent EKF SLAM Estimators**," *IJRR*, 2010.

Contestant 1: Sliding Window Filter

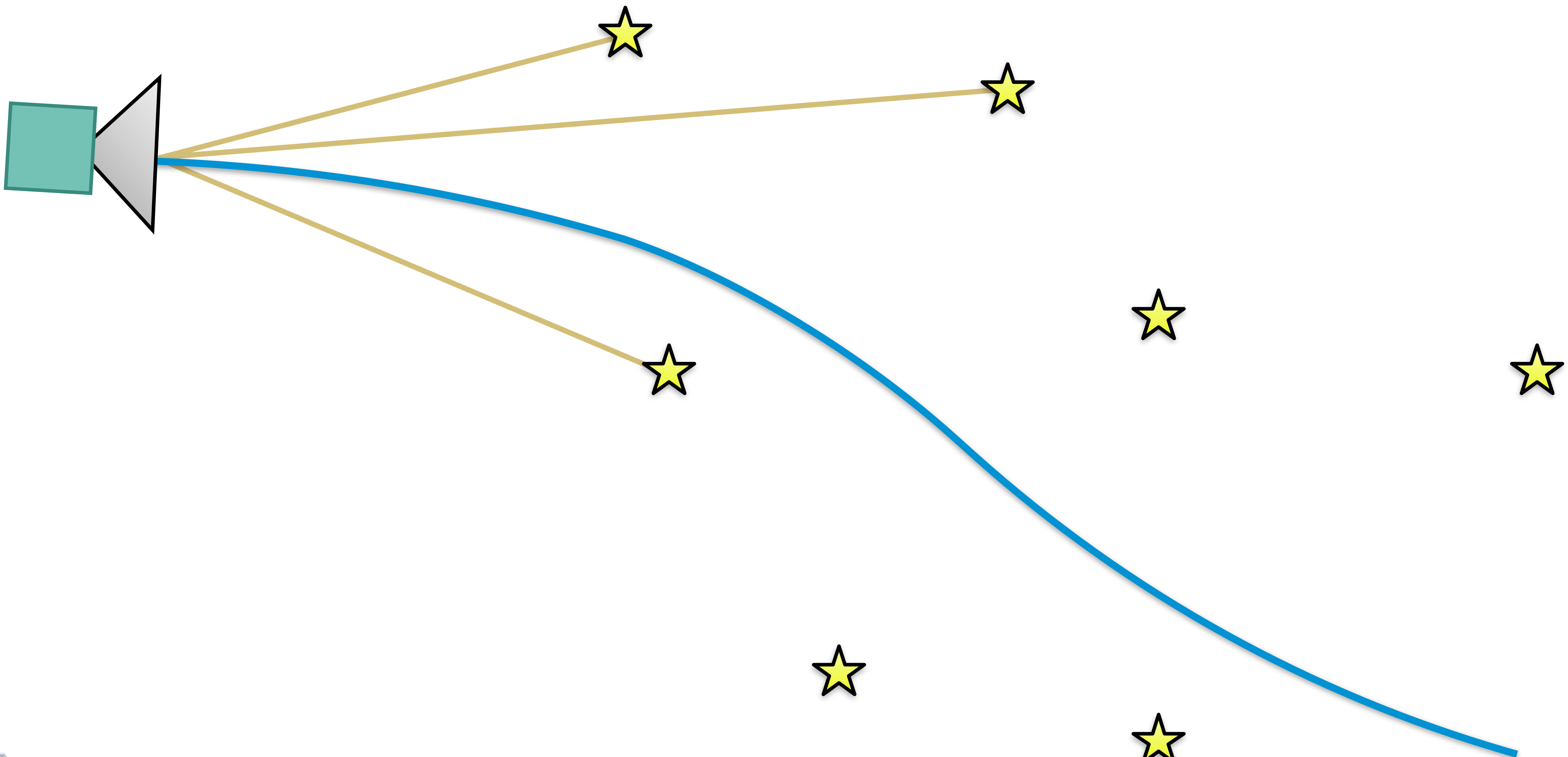
Given an initial guess of vehicle poses and feature positions in a fixed window, the SWF jointly optimizes poses and features in the window in a single batch operation.



Institute for Aerospace Studies
UNIVERSITY OF TORONTO

Contestant 1: Sliding Window Filter

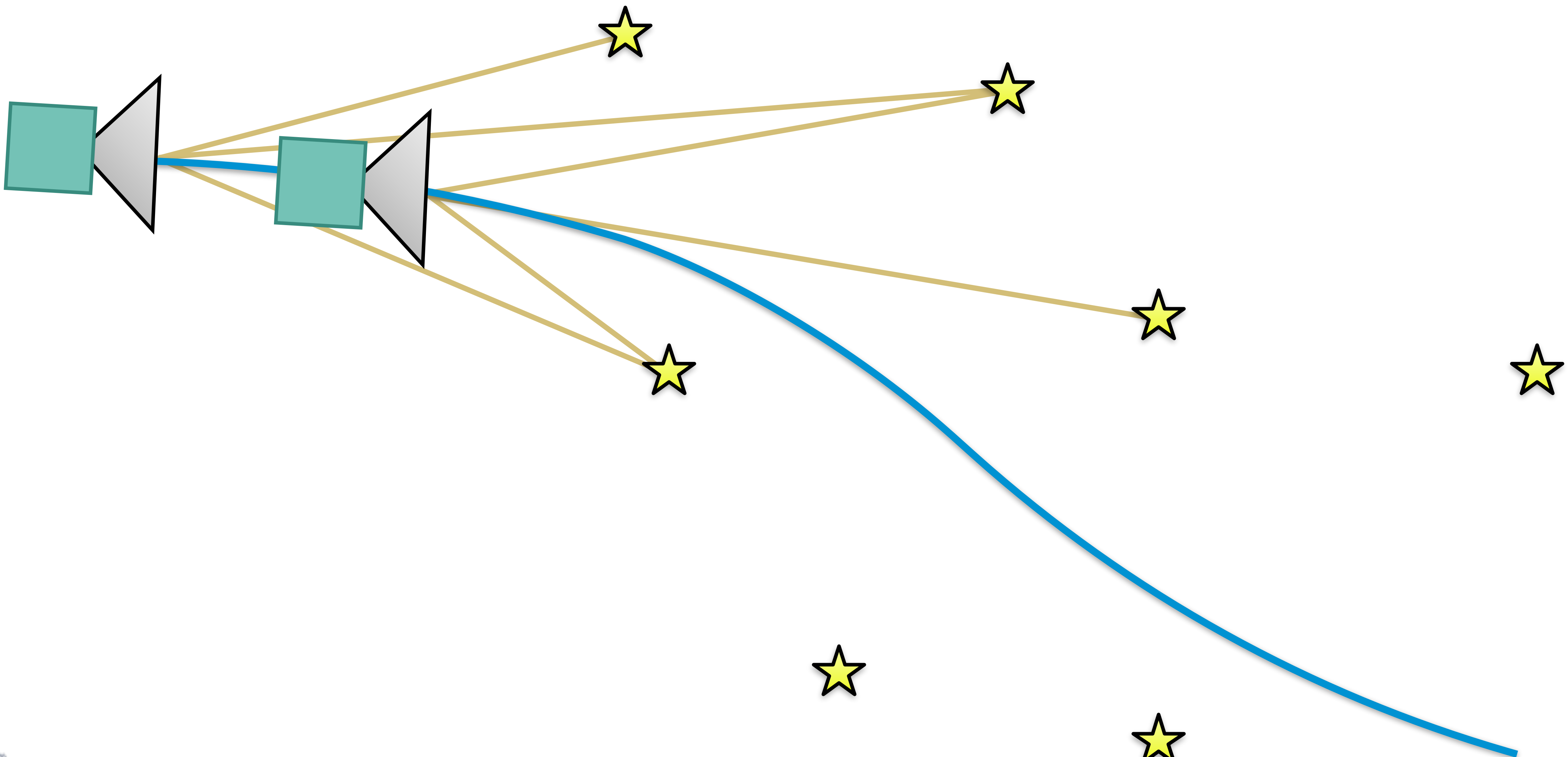
Given an initial guess of vehicle poses and feature positions in a fixed window, the SWF jointly optimizes poses and features in the window in a single batch operation.



Institute for Aerospace Studies
UNIVERSITY OF TORONTO

Contestant 1: Sliding Window Filter

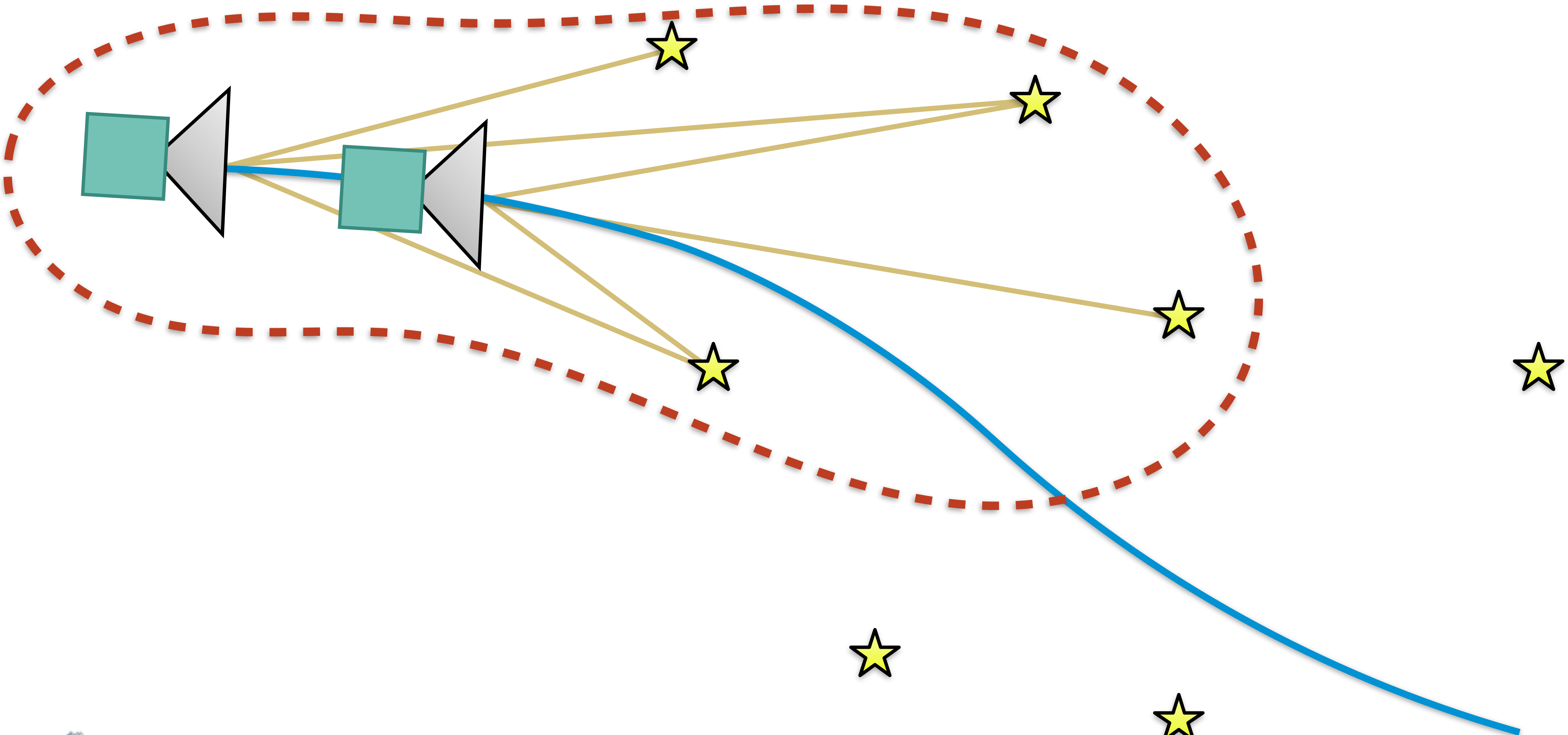
Given an initial guess of vehicle poses and feature positions in a fixed window, the SWF jointly optimizes poses and features in the window in a single batch operation.



Institute for Aerospace Studies
UNIVERSITY OF TORONTO

Contestant 1: Sliding Window Filter

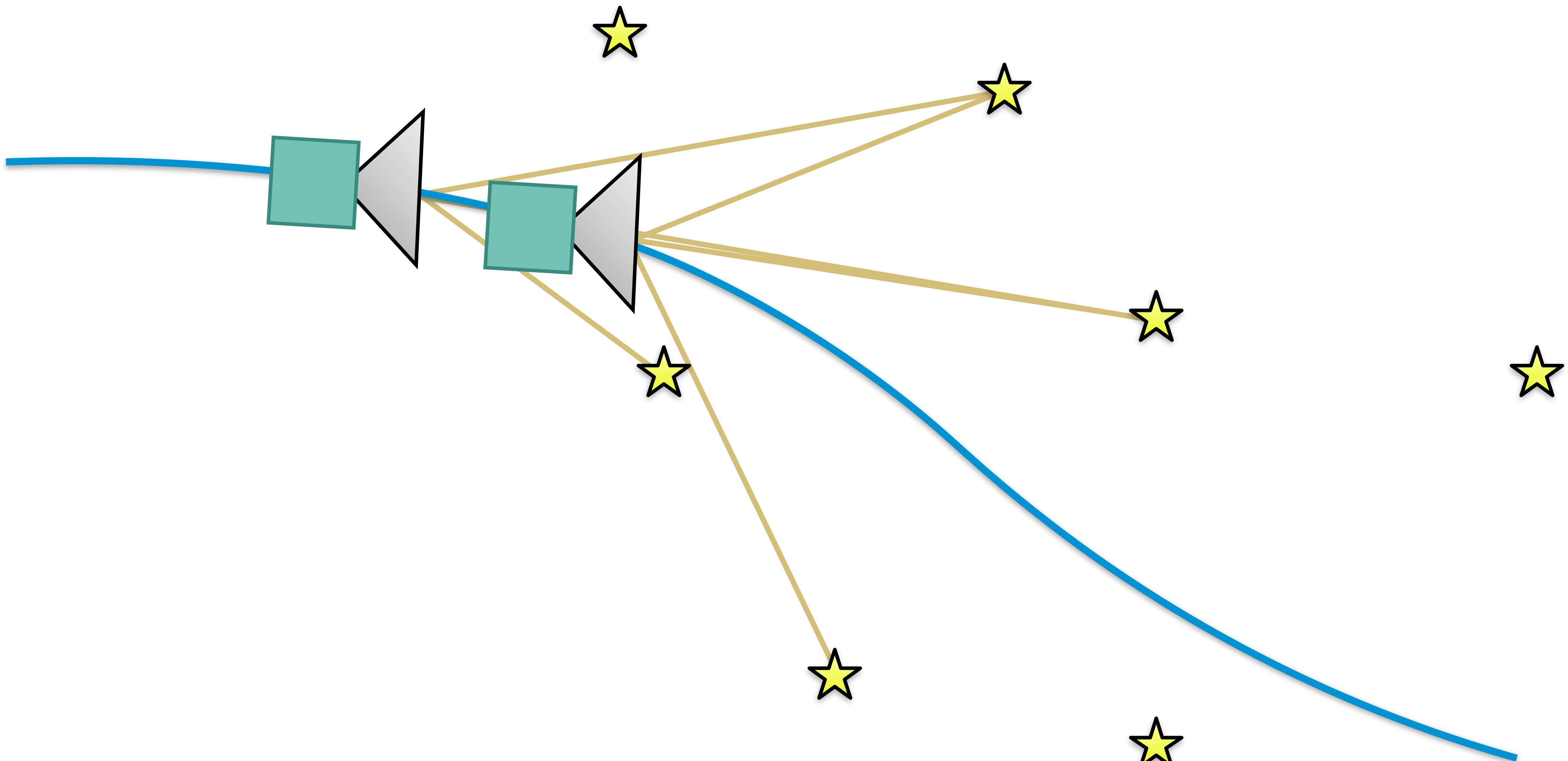
Given an initial guess of vehicle poses and feature positions in a fixed window, the SWF jointly optimizes poses and features in the window in a single batch operation.



Institute for Aerospace Studies
UNIVERSITY OF TORONTO

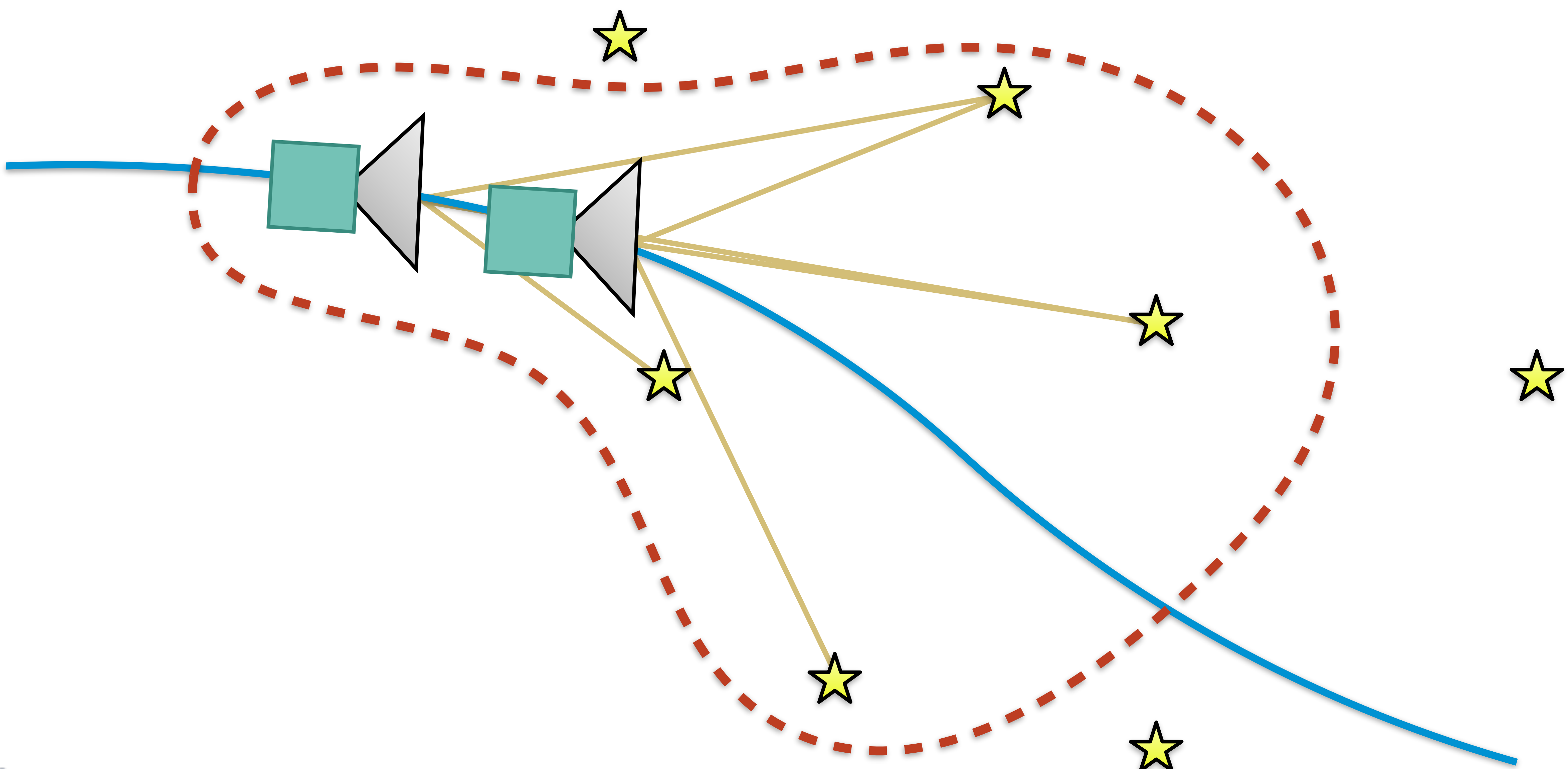
Contestant 1: Sliding Window Filter

Given an initial guess of vehicle poses and feature positions in a fixed window, the SWF jointly optimizes poses and features in the window in a single batch operation.



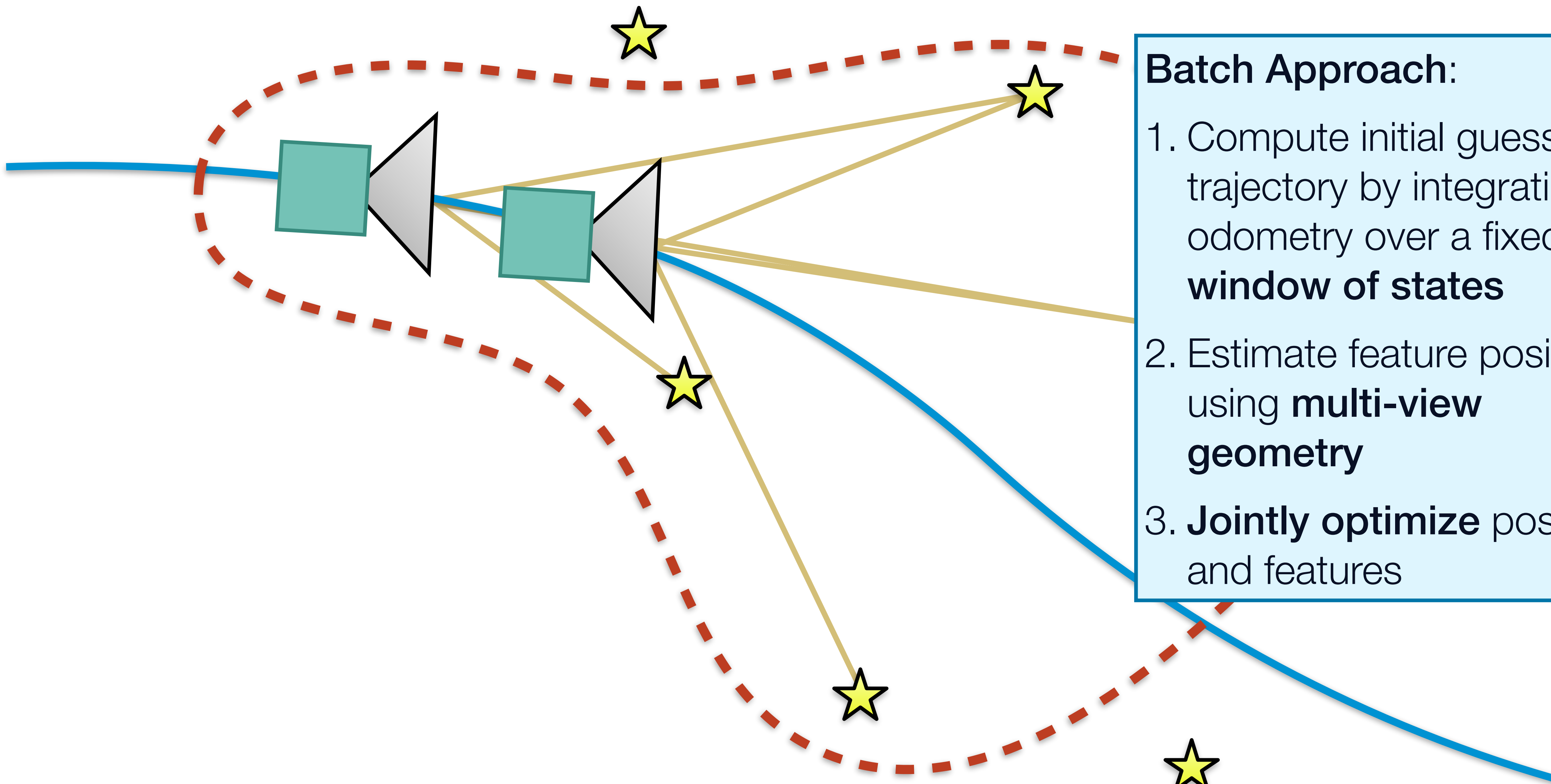
Contestant 1: Sliding Window Filter

Given an initial guess of vehicle poses and feature positions in a fixed window, the SWF jointly optimizes poses and features in the window in a single batch operation.



Contestant 1: Sliding Window Filter

Given an initial guess of vehicle poses and feature positions in a fixed window, the SWF jointly optimizes poses and features in the window in a single batch operation.



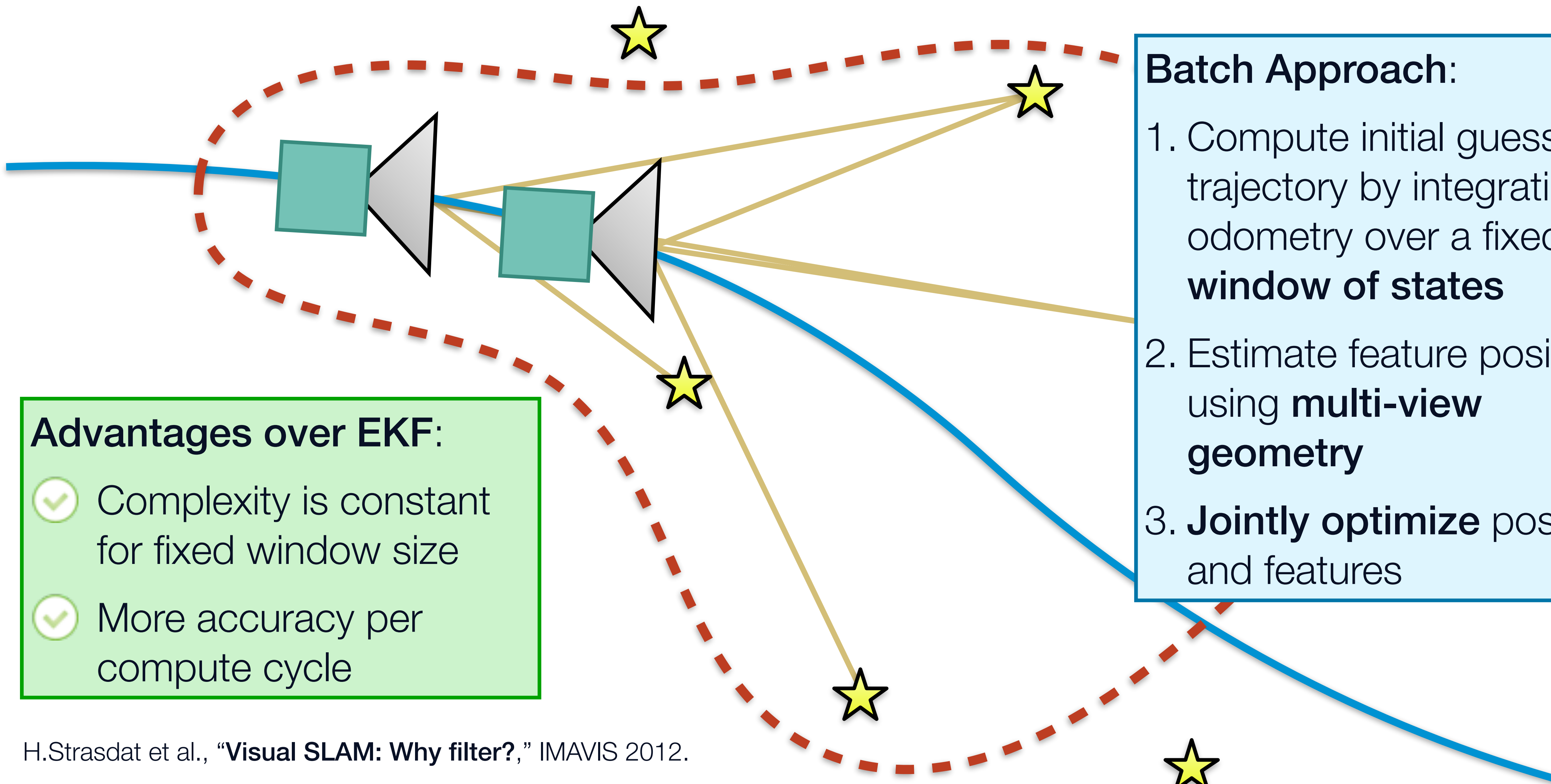
Batch Approach:

1. Compute initial guess for trajectory by integrating odometry over a fixed **window of states**
2. Estimate feature positions using **multi-view geometry**
3. **Jointly optimize** poses and features

Institute for Aerospace Studies
UNIVERSITY OF TORONTO

Contestant 1: Sliding Window Filter

Given an initial guess of vehicle poses and feature positions in a fixed window, the SWF jointly optimizes poses and features in the window in a single batch operation.



Batch Approach:

1. Compute initial guess for trajectory by integrating odometry over a fixed **window of states**
2. Estimate feature positions using **multi-view geometry**
3. **Jointly optimize** poses and features

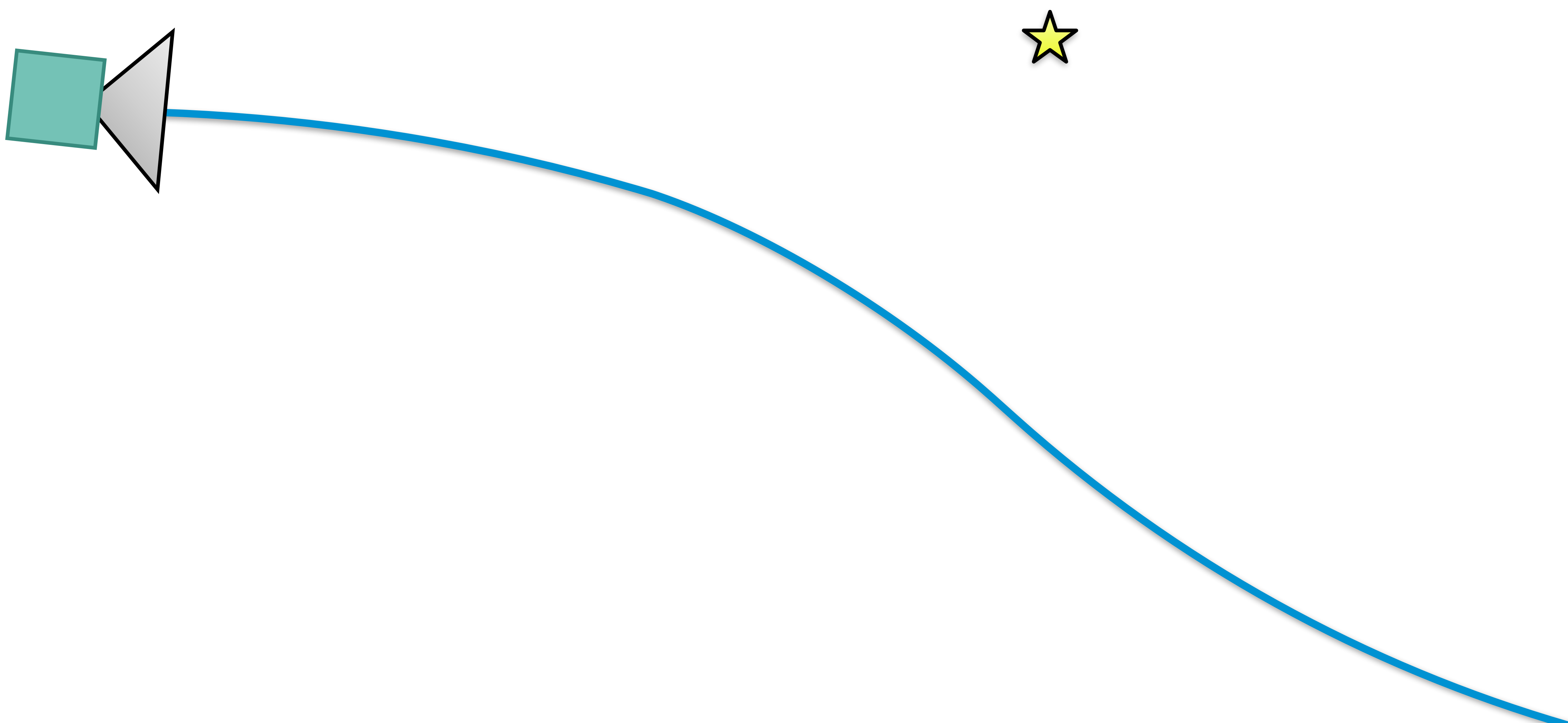
Advantages over EKF:

- Complexity is constant for fixed window size
- More accuracy per compute cycle

H.Strasdat et al., “**Visual SLAM: Why filter?**,” IMAVIS 2012.

Contestant 2: Multi-State Constraint Kalman Filter

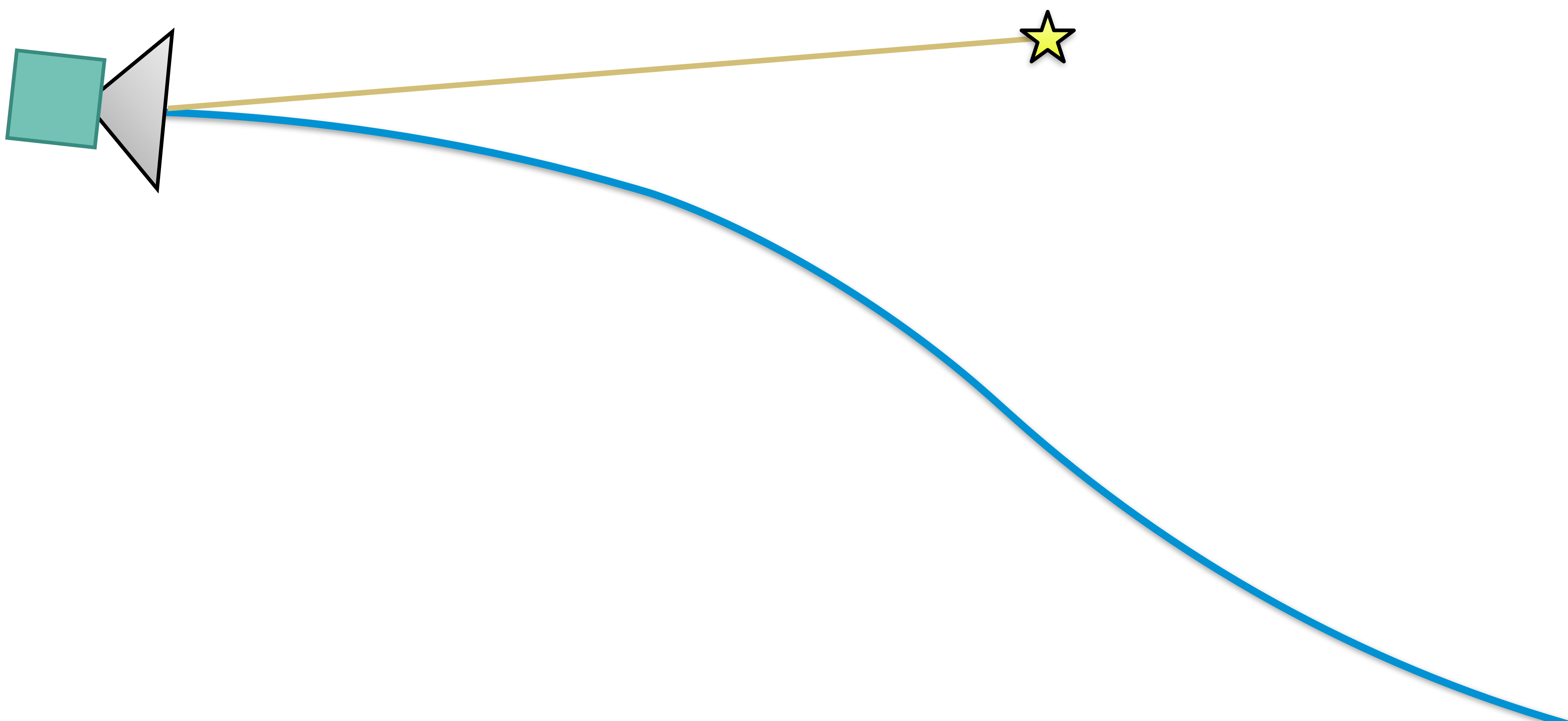
MSCKF dead-reckons the vehicle state using interoceptive (IMU) measurements, just like the EKF, but treats exteroceptive (camera) measurements like the SWF.



A. I. Mourikis and S. I. Roumeliotis, "**A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation**," ICRA, 2007.

Contestant 2: Multi-State Constraint Kalman Filter

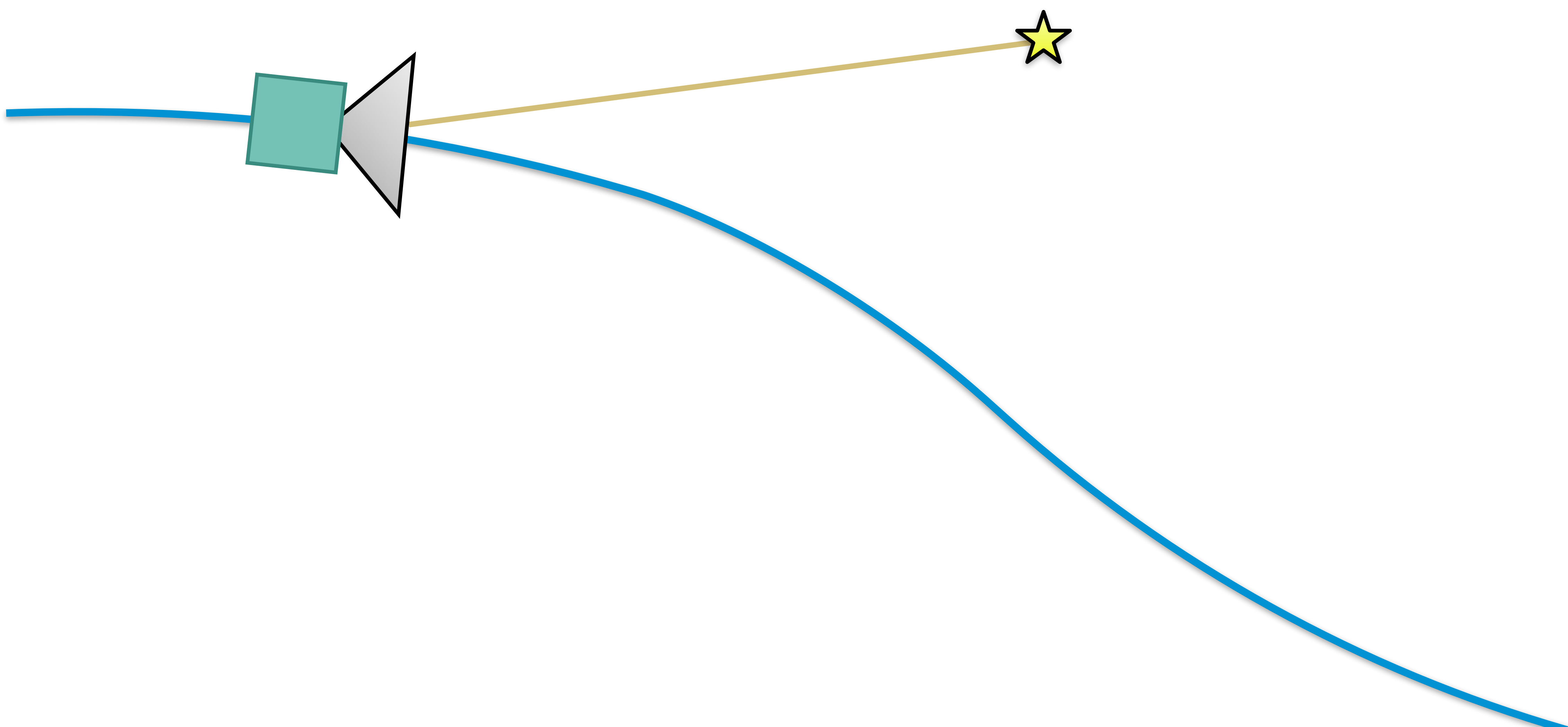
MSCKF dead-reckons the vehicle state using interoceptive (IMU) measurements, just like the EKF, but treats exteroceptive (camera) measurements like the SWF.



A. I. Mourikis and S. I. Roumeliotis, "**A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation**," ICRA, 2007.

Contestant 2: Multi-State Constraint Kalman Filter

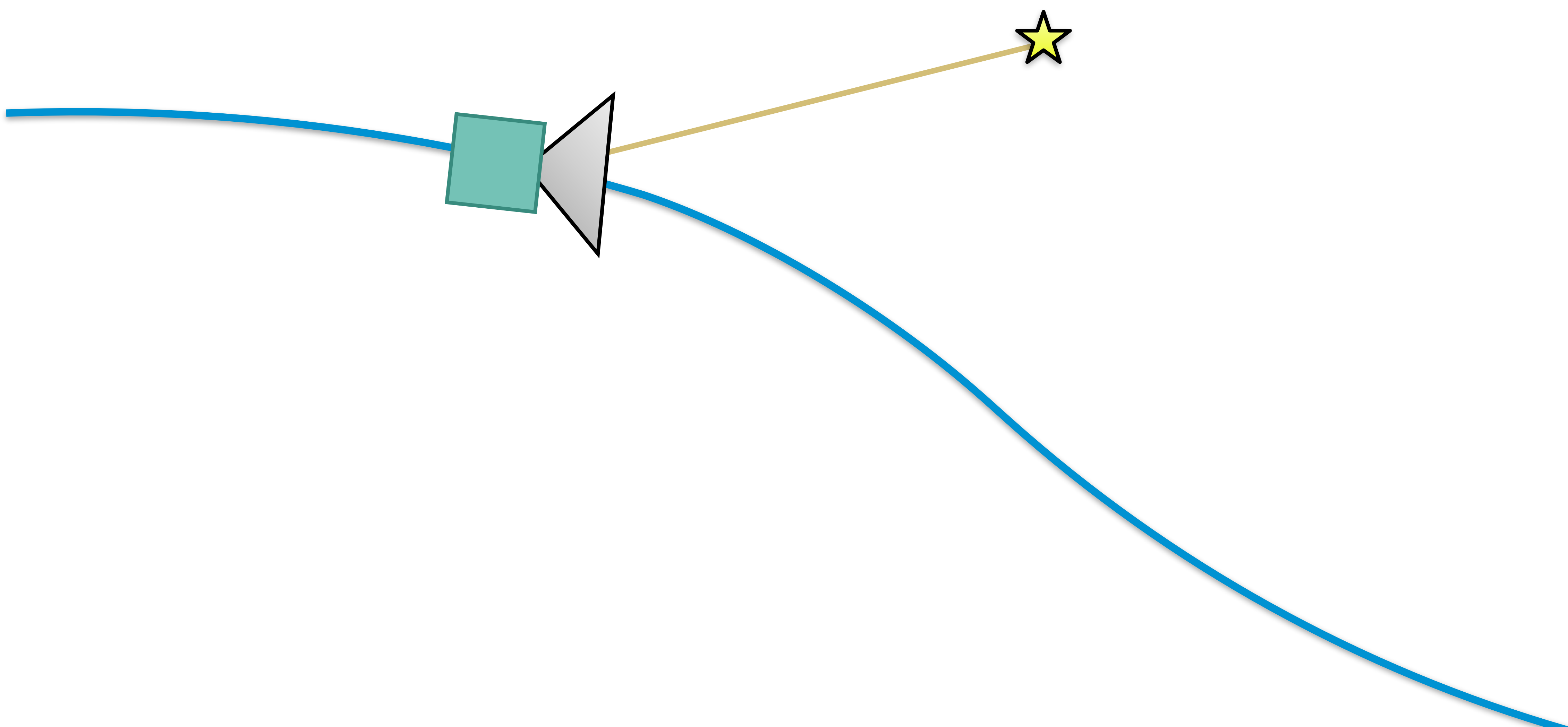
MSCKF dead-reckons the vehicle state using interoceptive (IMU) measurements, just like the EKF, but treats exteroceptive (camera) measurements like the SWF.



A. I. Mourikis and S. I. Roumeliotis, "**A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation**," ICRA, 2007.

Contestant 2: Multi-State Constraint Kalman Filter

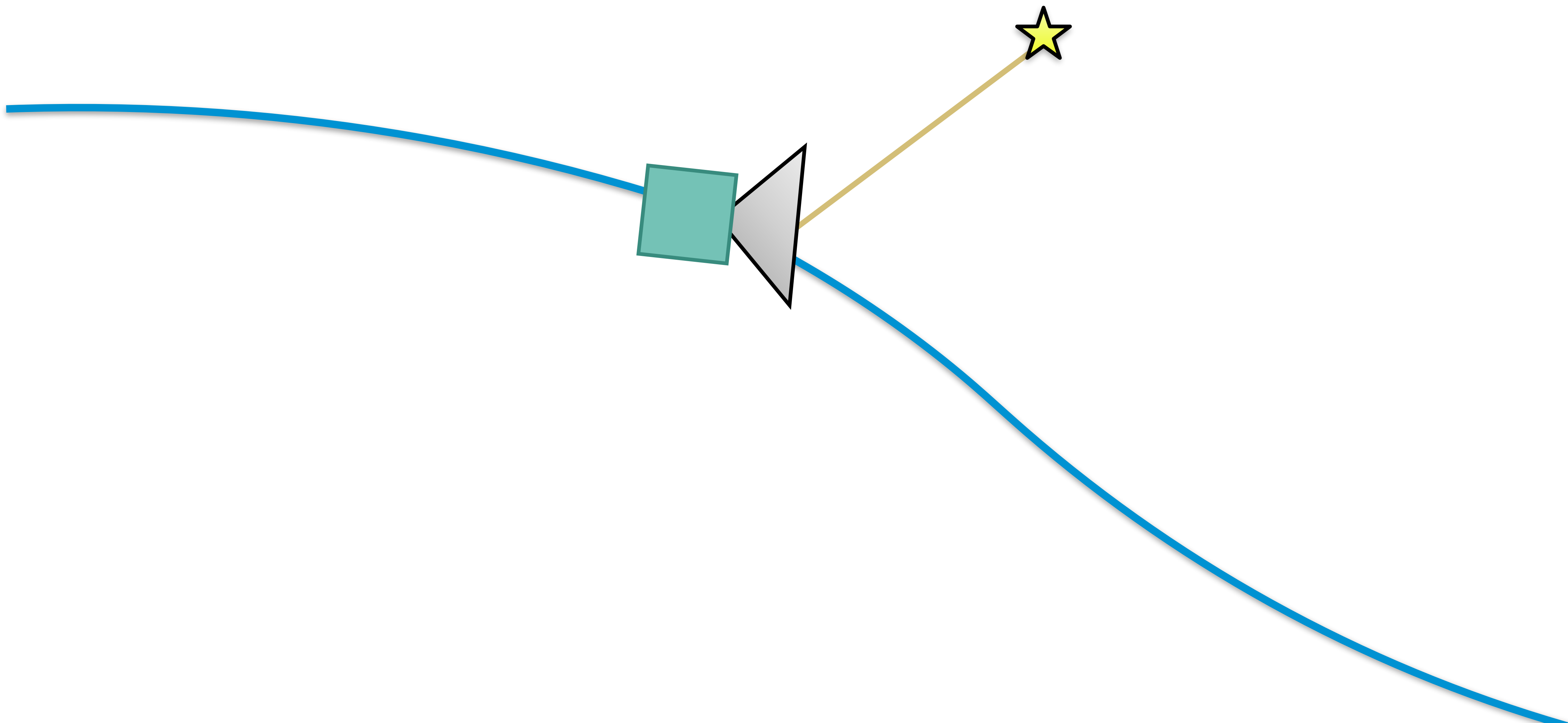
MSCKF dead-reckons the vehicle state using interoceptive (IMU) measurements, just like the EKF, but treats exteroceptive (camera) measurements like the SWF.



A. I. Mourikis and S. I. Roumeliotis, "**A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation**," ICRA, 2007.

Contestant 2: Multi-State Constraint Kalman Filter

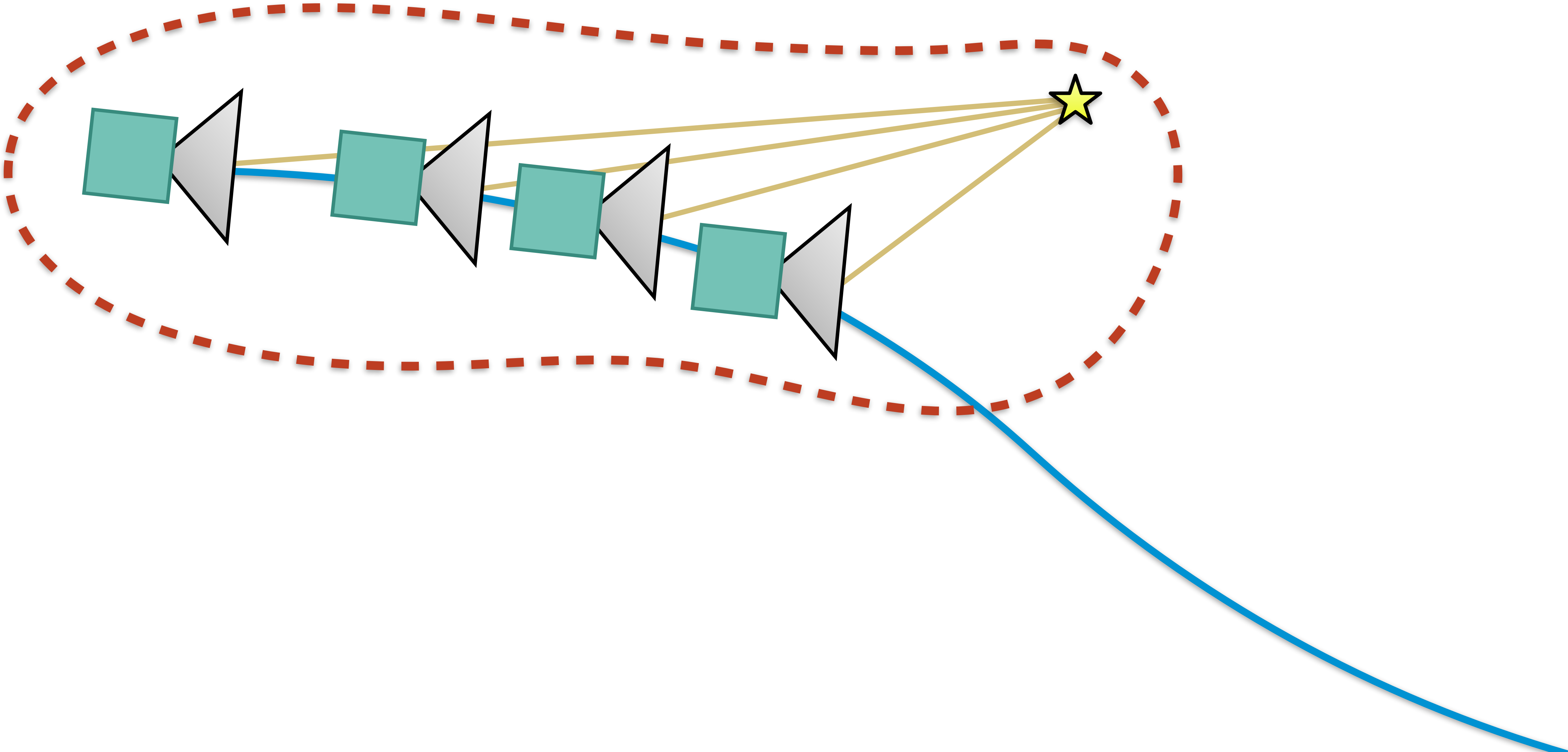
MSCKF dead-reckons the vehicle state using interoceptive (IMU) measurements, just like the EKF, but treats exteroceptive (camera) measurements like the SWF.



A. I. Mourikis and S. I. Roumeliotis, "**A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation**," ICRA, 2007.

Contestant 2: Multi-State Constraint Kalman Filter

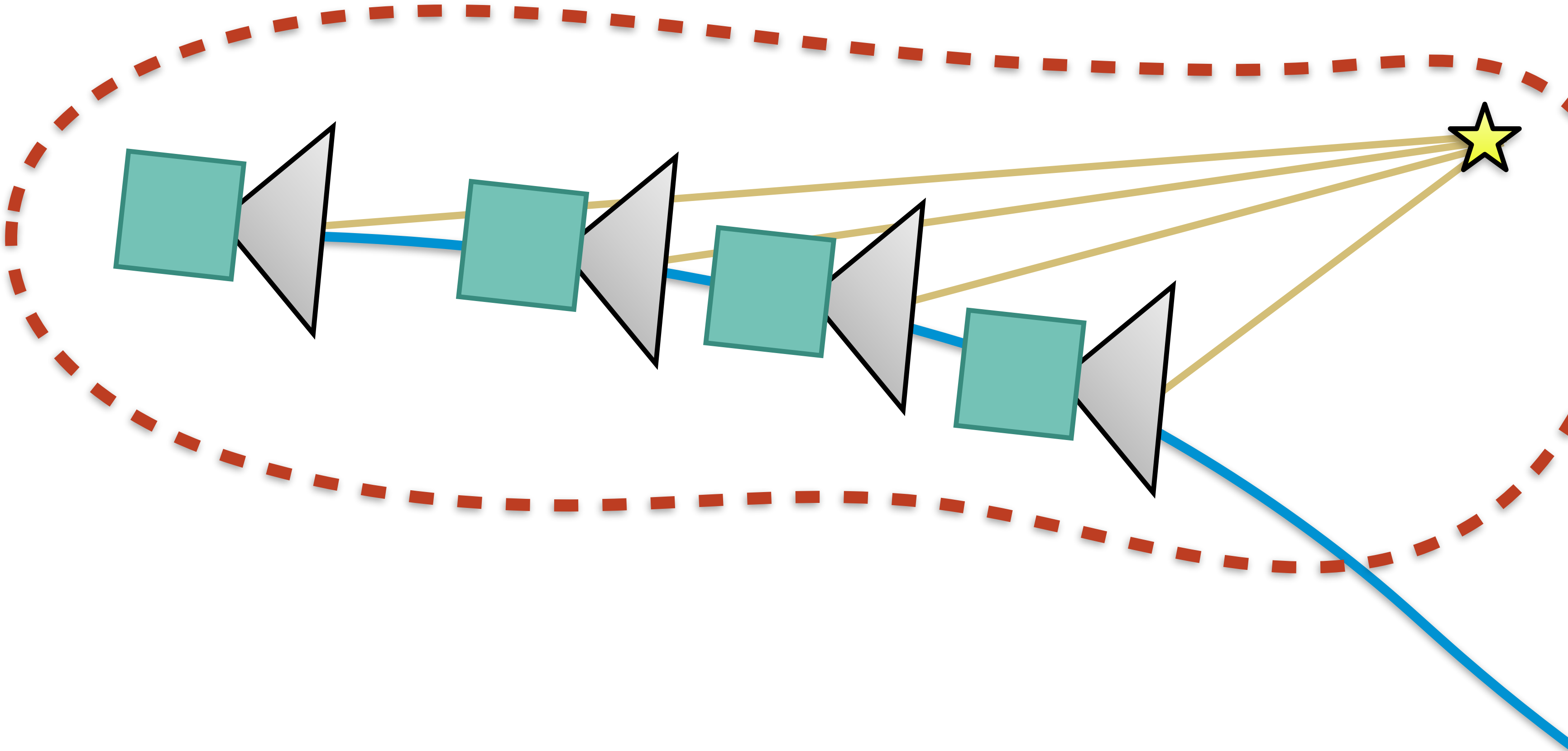
MSCKF dead-reckons the vehicle state using interoceptive (IMU) measurements, just like the EKF, but treats exteroceptive (camera) measurements like the SWF.



A. I. Mourikis and S. I. Roumeliotis, "**A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation**," ICRA, 2007.

Contestant 2: Multi-State Constraint Kalman Filter

MSCKF dead-reckons the vehicle state using interoceptive (IMU) measurements, just like the EKF, but treats exteroceptive (camera) measurements like the SWF.



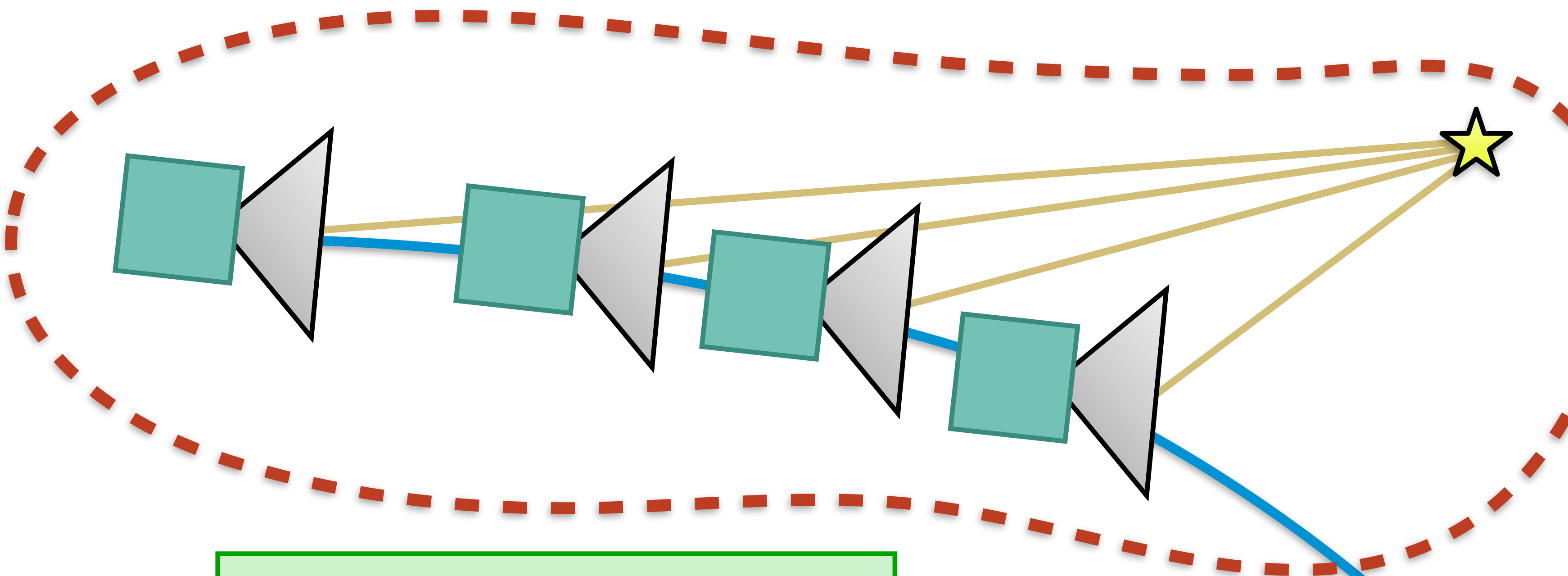
Hybrid Approach:

1. **Batch component** estimates feature position using Gauss-Newton
2. **Filter component** uses combined observations as one exteroceptive update
3. Repeat for all feature tracks over a **window of poses**

A. I. Mourikis and S. I. Roumeliotis, "**A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation**," ICRA, 2007.

Contestant 2: Multi-State Constraint Kalman Filter

MSCKF dead-reckons the vehicle state using interoceptive (IMU) measurements, just like the EKF, but treats exteroceptive (camera) measurements like the SWF.



Hybrid Approach:

1. **Batch component** estimates feature position using Gauss-Newton
2. **Filter component** uses combined observations as one exteroceptive update
3. Repeat for all feature tracks over a **window of poses**

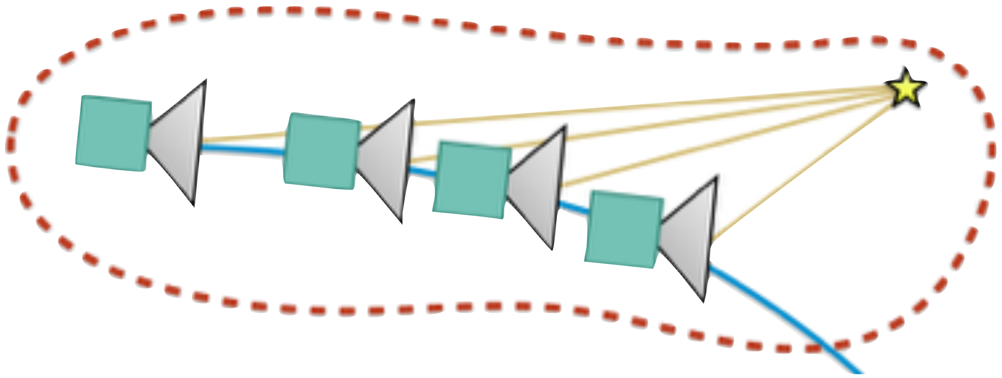
Advantages over EKF:

- Complexity is linear in number of features
- Each constraint affects multiple states, not just the most recent one

A. I. Mourikis and S. I. Roumeliotis, “**A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation**,” ICRA, 2007.

MSCKF: Null Space Trick

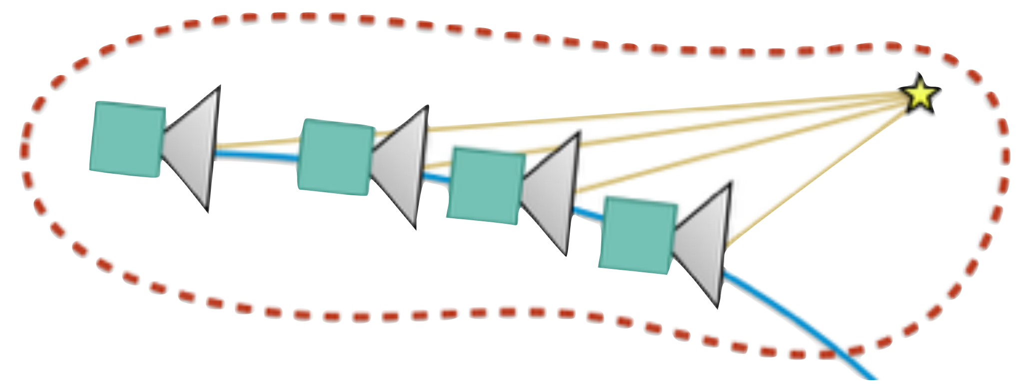
EKF: Many features constrain one state.



MSCKF: One feature constrains many states.

MSCKF: Null Space Trick

EKF: Many features constrain one state.



MSCKF: One feature constrains many states.

Measurement errors: $\mathbf{r}_i^{(j)} := \mathbf{z}_i^{(j)} - \hat{\mathbf{z}}_i^{(j)}$

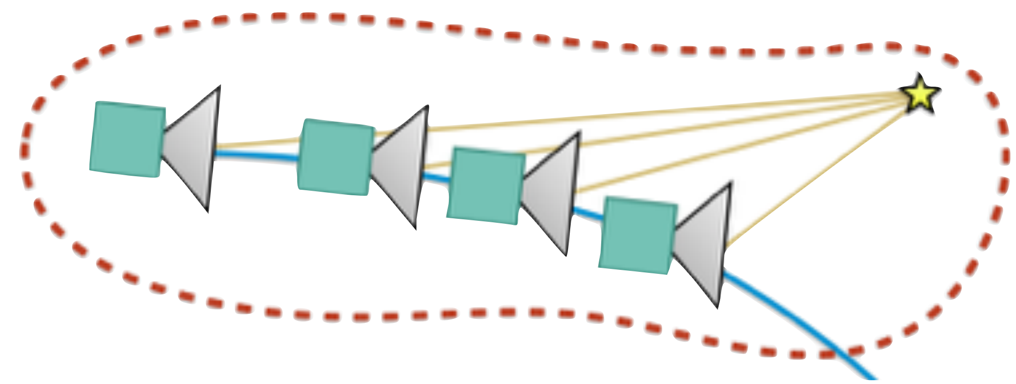
Feature position errors correlated with state!

Stacked and linearized… $\mathbf{r}^{(j)} = \mathbf{z}^{(j)} - \hat{\mathbf{z}}^{(j)} \simeq \mathbf{H}_{\mathbf{x}}^{(j)}\tilde{\mathbf{x}} + \mathbf{H}_f^{(j)}\tilde{\mathbf{p}}_G^{f_j G} + \mathbf{n}^{(j)}$

Institute for Aerospace Studies
UNIVERSITY OF TORONTO

MSCKF: Null Space Trick

EKF: Many features constrain one state.



MSCKF: One feature constrains many states.

Measurement errors: $\mathbf{r}_i^{(j)} := \mathbf{z}_i^{(j)} - \hat{\mathbf{z}}_i^{(j)}$

Feature position errors correlated with state!

Stacked and linearized… $\mathbf{r}^{(j)} = \mathbf{z}^{(j)} - \hat{\mathbf{z}}^{(j)} \simeq \mathbf{H}_{\mathbf{x}}^{(j)}\tilde{\mathbf{x}} + \mathbf{H}_f^{(j)}\,{}^{G}\tilde{\mathbf{p}}_{f_j} + \mathbf{n}^{(j)}$

Not correlated with state!

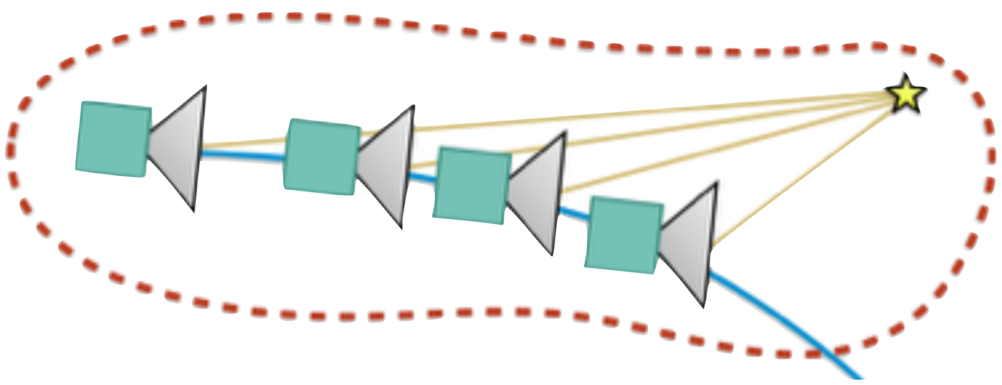
Project into nullspace of $\mathbf{H}_f^{(j)}$: $\mathbf{r}_o^{(j)} := \mathbf{A}^T\mathbf{r}^{(j)} \simeq \mathbf{A}^T\mathbf{H}_{\mathbf{x}}^{(j)}\tilde{\mathbf{x}} + \mathbf{0} + \mathbf{A}^T\mathbf{n}^{(j)}$

$\mathbf{A} = \mathrm{Null}(\mathbf{H}_f^{(j)})$ $\qquad =: \mathbf{H}_o^{(j)}\tilde{\mathbf{x}} + \mathbf{n}_o^{(j)}$

Stacked… $\mathbf{r}_o = \mathbf{H}_o\tilde{\mathbf{x}} + \mathbf{n}_o$

MSCKF: Null Space Trick

EKF: Many features constrain one state.



MSCKF: One feature constrains many states.

Measurement errors: $\mathbf{r}_i^{(j)} := \mathbf{z}_i^{(j)} - \hat{\mathbf{z}}_i^{(j)}$

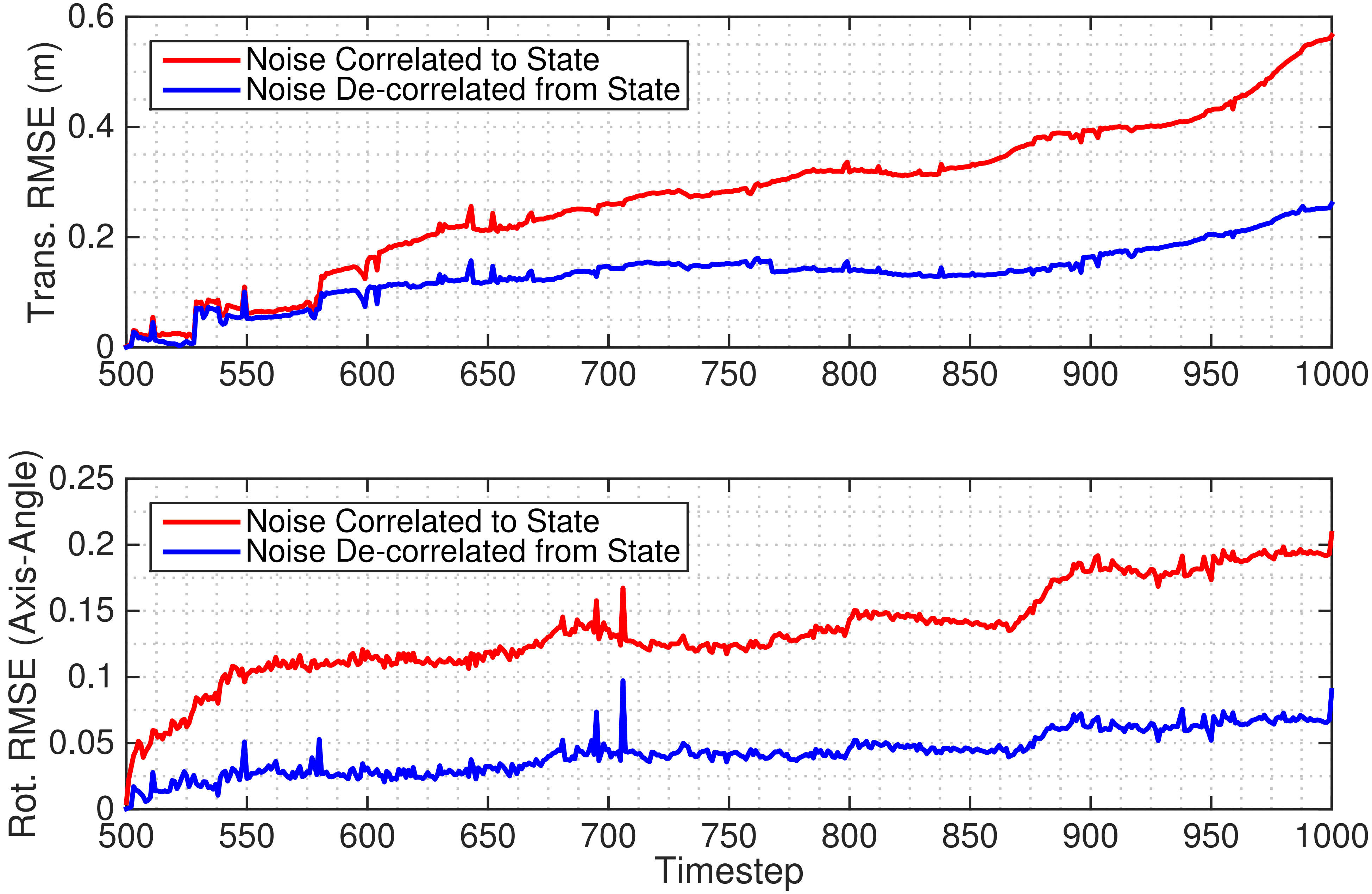
Feature position errors correlated with state!

Stacked and linearized… $\mathbf{r}^{(j)} = \mathbf{z}^{(j)} - \hat{\mathbf{z}}^{(j)} \simeq \mathbf{H}_{\mathbf{x}}^{(j)}\widetilde{\mathbf{x}} + \mathbf{H}_f^{(j)}\widetilde{\mathbf{p}}_G^{f_j G} + \mathbf{n}^{(j)}$

Is this null space projection really necessary?

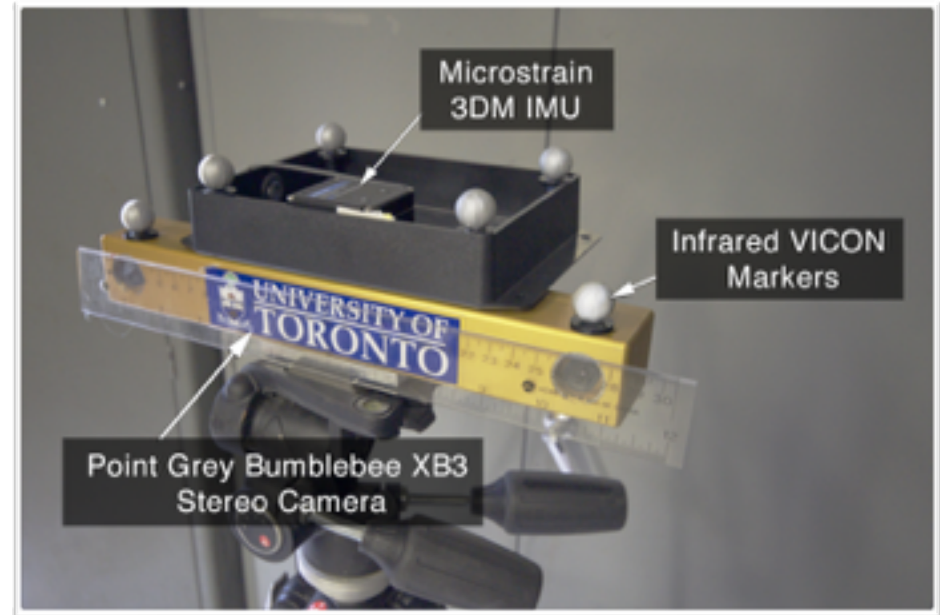
Institute for Aerospace Studies
UNIVERSITY OF TORONTO

MSCKF: Null Space Trick



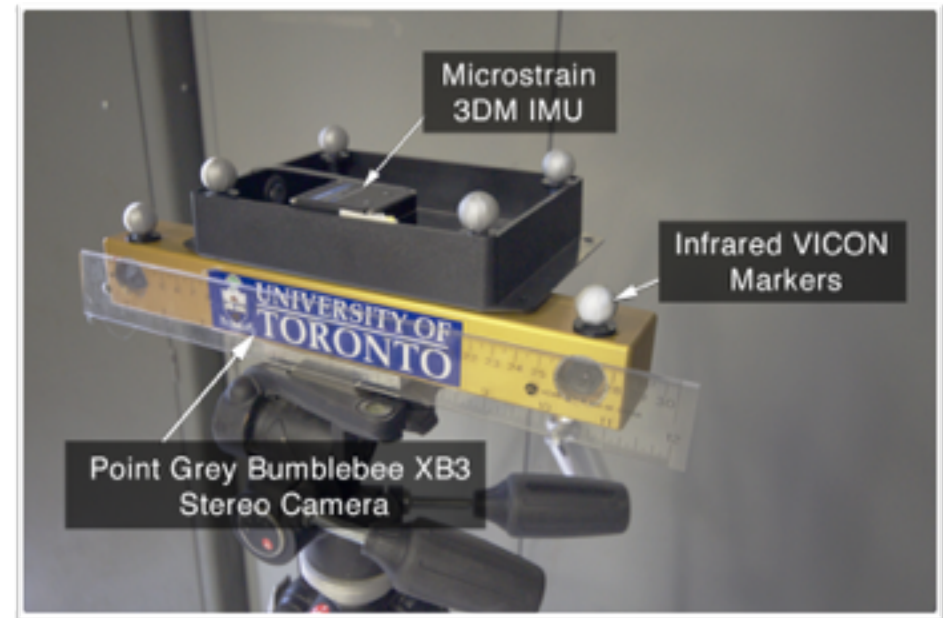
Institute for Aerospace Studies
UNIVERSITY OF TORONTO

Experiment 1: Starry Night Dataset



- ☑ Perfect data association
- ☑ Ground truth for landmark positions
- ☑ Pre-integrated IMU measurements

Experiment 1: Starry Night Dataset

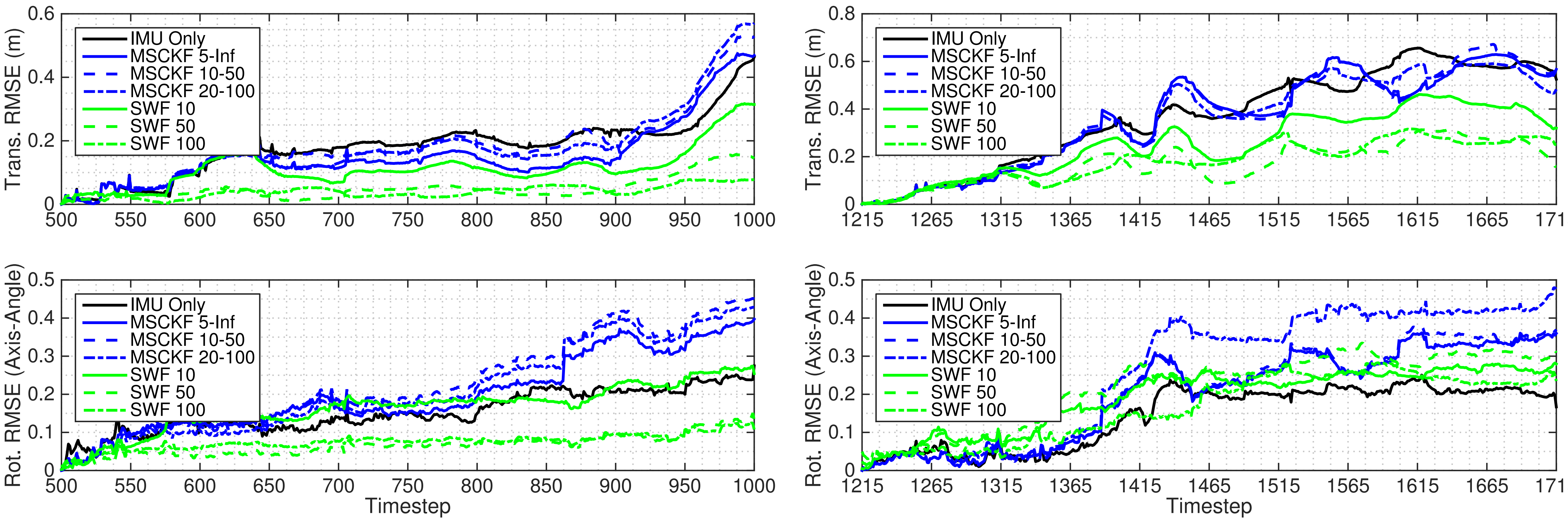


- ☑ Perfect data association
- ☑ Ground truth for landmark positions
- ☑ Pre-integrated IMU measurements

Experiment 1.1: Window Size Comparison

We investigated the sensitivity of estimation accuracy to window size.

MSCKF SWF



Many visible features

Few visible features

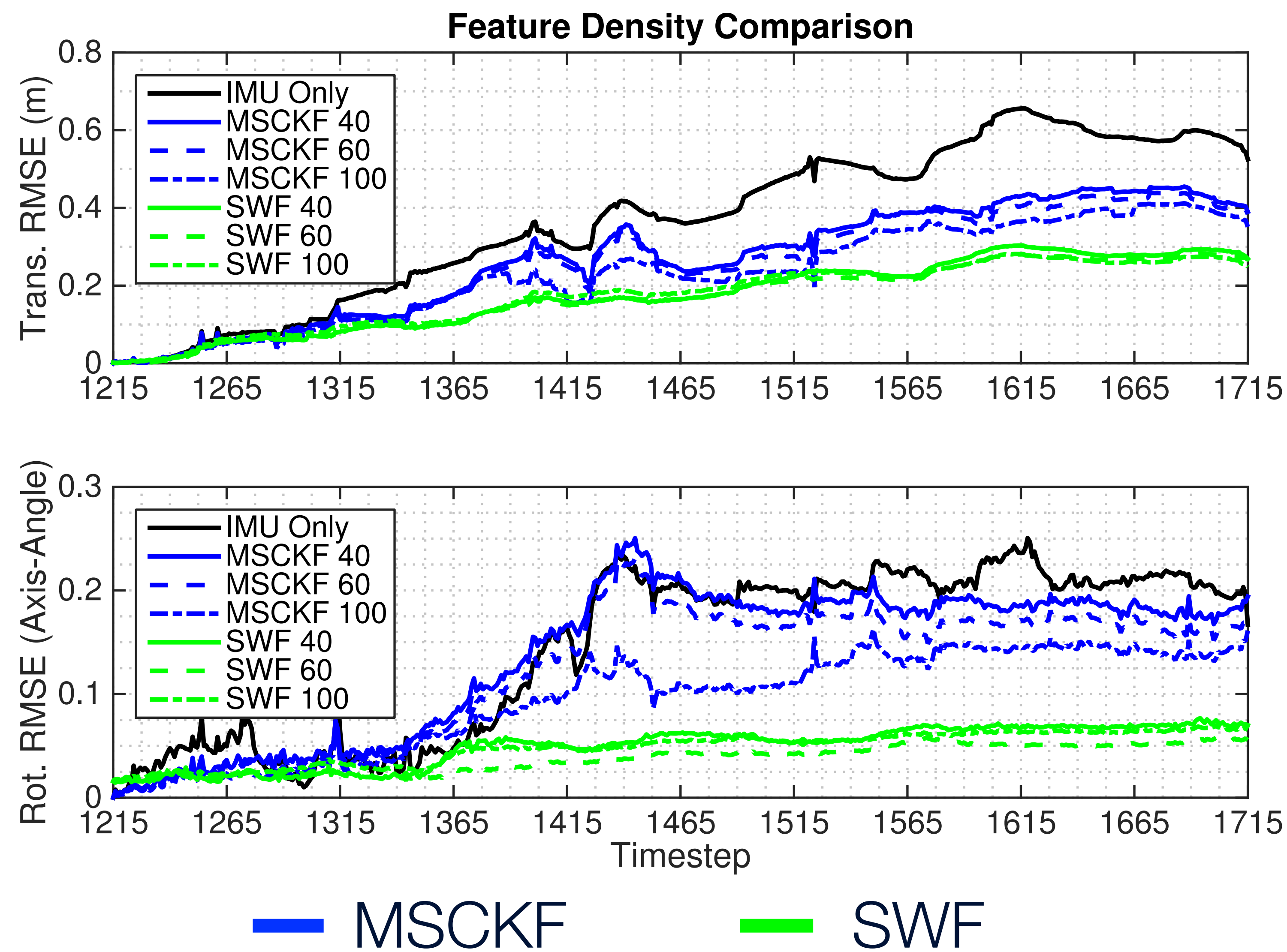
Institute for Aerospace Studies
UNIVERSITY OF TORONTO

The Battle for Filter Supremacy: A Comparative Study of the MSCKF and the SWF

Lee Clement, Valentin Peretroukhin, Jacob Lambert, and Jonathan Kelly

Experiment 1.2: Feature Density Comparison

We added synthetic features to the dataset to investigate the sensitivity of estimation accuracy to feature density.



Institute for Aerospace Studies
UNIVERSITY OF TORONTO

Experiment 2: KITTI Dataset

- ☑ Realistic urban driving
- ☑ High quality IMU data
- ☑ Synchronized measurements

A. Geiger et al. "**Vision meets robotics: The KITTI dataset**," IJRR 2013. http://www.cvlibs.net/datasets/kitti/

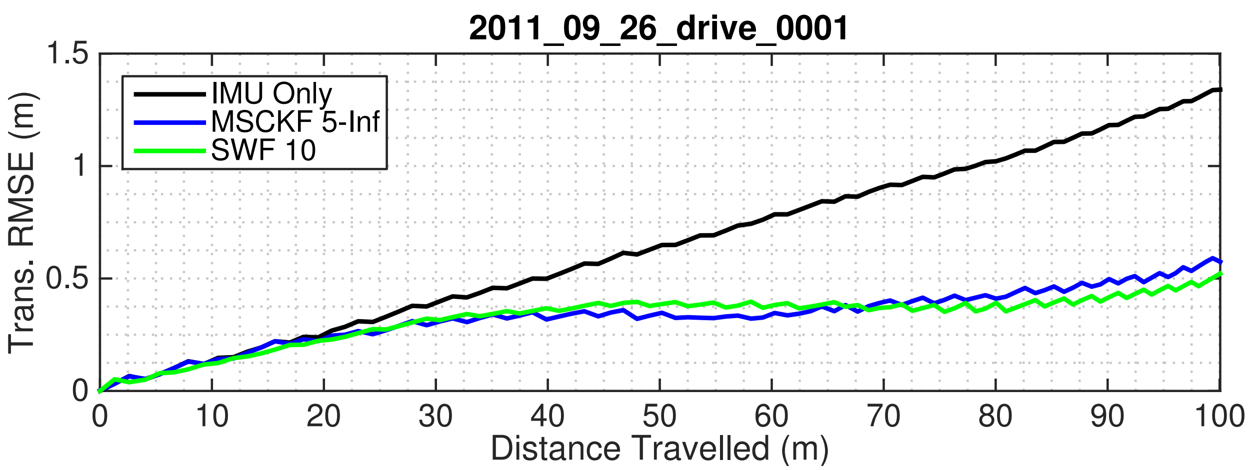
Institute for Aerospace Studies
UNIVERSITY OF TORONTO

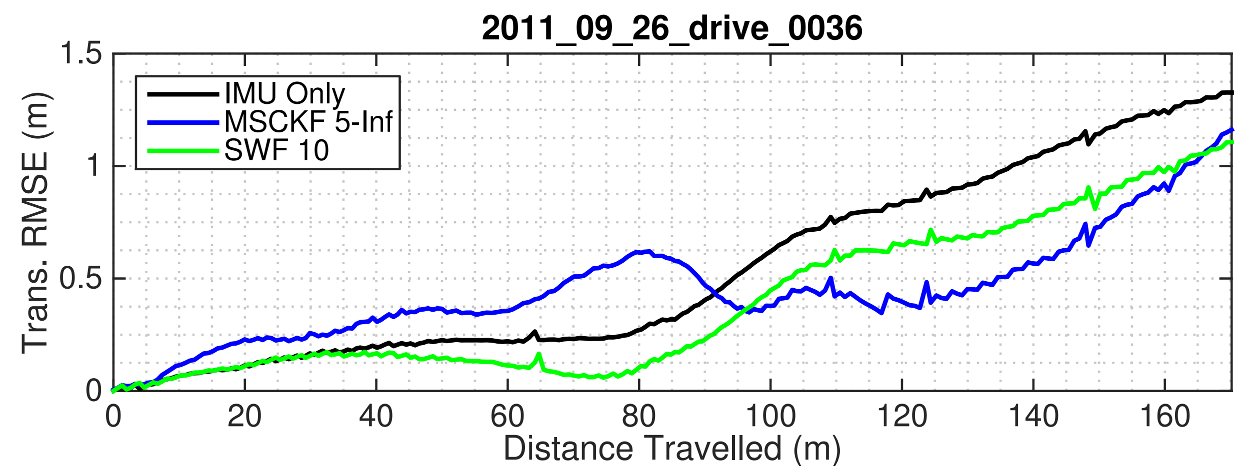
Experiment 2: KITTI Dataset

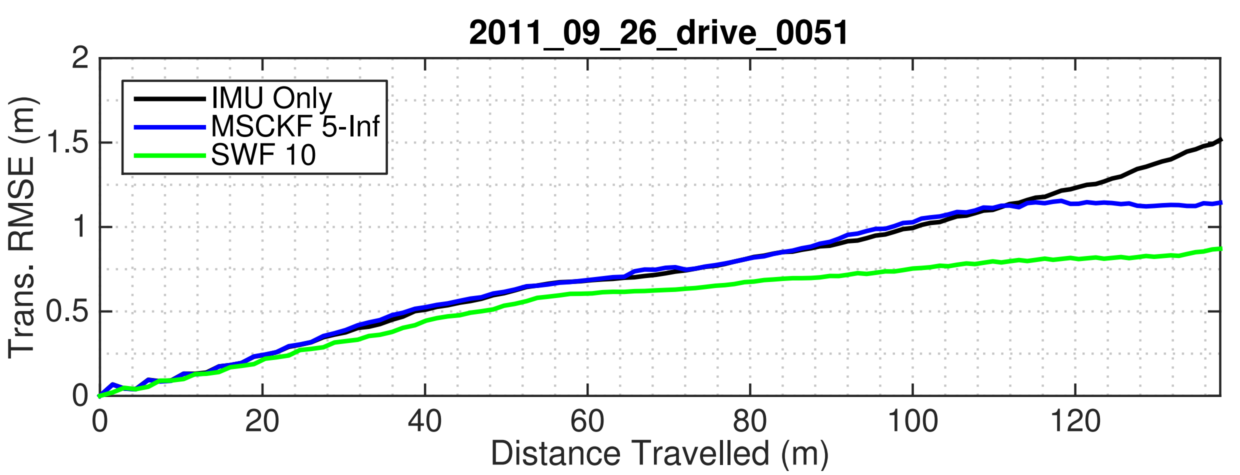
- ☑ Realistic urban driving
- ☑ High quality IMU data
- ☑ Synchronized measurements

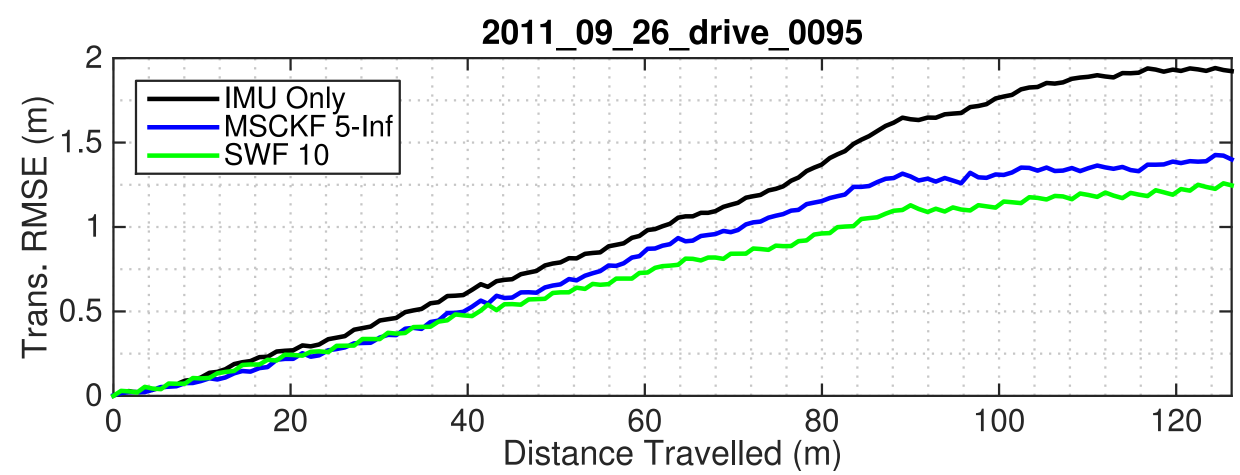
A. Geiger et al. "**Vision meets robotics: The KITTI dataset**," IJRR 2013. http://www.cvlibs.net/datasets/kitti/

Experiment 2: KITTI Dataset









Experiments: Summary

Starry Night

		Feature Count		
		40	60	100
IMU Only	Trans. ARMSE	0.3679	0.3679	0.3679
	Rot. ARMSE	0.1452	0.1452	0.1452
	ANEES	0.2850	0.2850	0.2850
	Compute Time†	8.90 s	8.90 s	8.90 s
MSCKF (20-100)	Trans. ARMSE	0.2672	0.2550	0.2304
	Rot. ARMSE	0.1378	0.1247	0.0952
	ANEES	10.18	12.03	16.76
	Compute Time†	12.19 s	14.64 s	20.58 s
SWF (25)	Trans. ARMSE	0.1750	0.1687	0.1755
	Rot. ARMSE	0.0495	0.0377	0.0481
	ANEES	2280	2093	2013
	Compute Time†	114.3 s	175.9 s	245.3 s

† Running MATLAB 2014b on a MacBook Pro Retina (11,3) with a 2.3 GHz Intel Core i7 processor and 16 GB of DDR3L RAM.

KITTI

		KITTI Traverse			
		0001	0036	0051	0095
IMU Only	Trans. ARMSE	0.7197	0.5131	0.7834	1.039
	ANEES	0.1630	0.0092	0.1170	0.6254
MSCKF (5-Inf)	Trans. ARMSE	0.3492	0.4401	0.7530	0.8170
	ANEES	5.103	1.826	2.031	14.98
SWF (10)	Trans. ARMSE	0.3372	0.3778	0.5832	0.7196
	ANEES	358.3	703.2	1124	3767

Experiments: Summary

Starry Night

		Feature Count		
		40	60	100
IMU Only	Trans. ARMSE	0.3679	0.3679	0.3679
	Rot. ARMSE	0.1452	0.1452	0.1452
	ANEES	0.2850	0.2850	0.2850
	Compute Time†	8.90 s	8.90 s	8.90 s
MSCKF (20-100)	Trans. ARMSE	0.2672	0.2550	0.2304
	Rot. ARMSE	0.1378	0.1247	0.0952
	ANEES	10.18	12.03	16.76
	Compute Time†	12.19 s	14.64 s	20.58 s
SWF (25)	Trans. ARMSE	0.1750	0.1687	0.1755
	Rot. ARMSE	0.0495	0.0377	0.0481
	ANEES	2280	2093	2013
	Compute Time†	114.3 s	175.9 s	245.3 s

† Running MATLAB 2014b on a MacBook Pro Retina (11,3) with a 2.3 GHz Intel Core i7 processor and 16 GB of DDR3L RAM.

KITTI

		KITTI Traverse			
		0001	0036	0051	0095
IMU Only	Trans. ARMSE	0.7197	0.5131	0.7834	1.039
	ANEES	0.1630	0.0092	0.1170	0.6254
MSCKF (5-Inf)	Trans. ARMSE	0.3492	0.4401	0.7530	0.8170
	ANEES	5.103	1.826	2.031	14.98
SWF (10)	Trans. ARMSE	0.3372	0.3778	0.5832	0.7196
	ANEES	358.3	703.2	1124	3767

Experiments: Summary

Starry Night

		Feature Count		
		40	60	100
IMU Only	Trans. ARMSE	0.3679	0.3679	0.3679
	Rot. ARMSE	0.1452	0.1452	0.1452
	ANEES	0.2850	0.2850	0.2850
	Compute Time†	8.90 s	8.90 s	8.90 s
MSCKF (20-100)	Trans. ARMSE	0.2672	0.2550	0.2304
	Rot. ARMSE	0.1378	0.1247	0.0952
	ANEES	10.18	12.03	16.76
	Compute Time†	12.19 s	14.64 s	20.58 s
SWF (25)	Trans. ARMSE	0.1750	0.1687	0.1755
	Rot. ARMSE	0.0495	0.0377	0.0481
	ANEES	2280	2093	2013
	Compute Time†	114.3 s	175.9 s	245.3 s

† Running MATLAB 2014b on a MacBook Pro Retina (11,3) with a 2.3 GHz Intel Core i7 processor and 16 GB of DDR3L RAM.

KITTI

		KITTI Traverse			
		0001	0036	0051	0095
IMU Only	Trans. ARMSE	0.7197	0.5131	0.7834	1.039
	ANEES	0.1630	0.0092	0.1170	0.6254
MSCKF (5-Inf)	Trans. ARMSE	0.3492	0.4401	0.7530	0.8170
	ANEES	5.103	1.826	2.031	14.98
SWF (10)	Trans. ARMSE	0.3372	0.3778	0.5832	0.7196
	ANEES	358.3	703.2	1124	3767

Institute for Aerospace Studies
UNIVERSITY OF TORONTO

Experiments: Summary

Starry Night

		Feature Count		
		40	60	100
IMU Only	Trans. ARMSE	0.3679	0.3679	0.3679
	Rot. ARMSE	0.1452	0.1452	0.1452
	ANEES	0.2850	0.2850	0.2850
	Compute Time†	8.90 s	8.90 s	8.90 s
MSCKF (20-100)	Trans. ARMSE	0.2672	0.2550	0.2304
	Rot. ARMSE	0.1378	0.1247	0.0952
	ANEES	10.18	12.03	16.76
	Compute Time†	12.19 s	14.64 s	20.58 s
SWF (25)	Trans. ARMSE	0.1750	0.1687	0.1755
	Rot. ARMSE	0.0495	0.0377	0.0481
	ANEES	2280	2093	2013
	Compute Time†	114.3 s	175.9 s	245.3 s

† Running MATLAB 2014b on a MacBook Pro Retina (11,3) with a 2.3 GHz Intel Core i7 processor and 16 GB of DDR3L RAM.

KITTI

		KITTI Traverse			
		0001	0036	0051	0095
IMU Only	Trans. ARMSE	0.7197	0.5131	0.7834	1.039
	ANEES	0.1630	0.0092	0.1170	0.6254
MSCKF (5-Inf)	Trans. ARMSE	0.3492	0.4401	0.7530	0.8170
	ANEES	5.103	1.826	2.031	14.98
SWF (10)	Trans. ARMSE	0.3372	0.3778	0.5832	0.7196
	ANEES	358.3	703.2	1124	3767

Institute for Aerospace Studies
UNIVERSITY OF TORONTO

The Battle for Filter Supremacy: Who Won?

	Winner	Comments
Accuracy	**SWF**	Accuracy of MSCKF is more sensitive to length and number of feature tracks.
Consistency (in basic form)	**MSCKF**	Consistency of SWF can be improved by marginalizing out old poses. (Sibley et al., 2010)
Compute time	**MSCKF**	MSCKF complexity scales linearly with number of features, SWF complexity scales cubically in general.
Sensitivity to tuning parameters	**SWF**	In our experience, MSCKF is very difficult to tune for optimal performance.

Thanks!
Questions?

Email: {lee.clement, v.peretroukhin, jacob.lambert}@mail.utoronto.ca, jkelly@utias.utoronto.ca

Web: http://starslab.ca

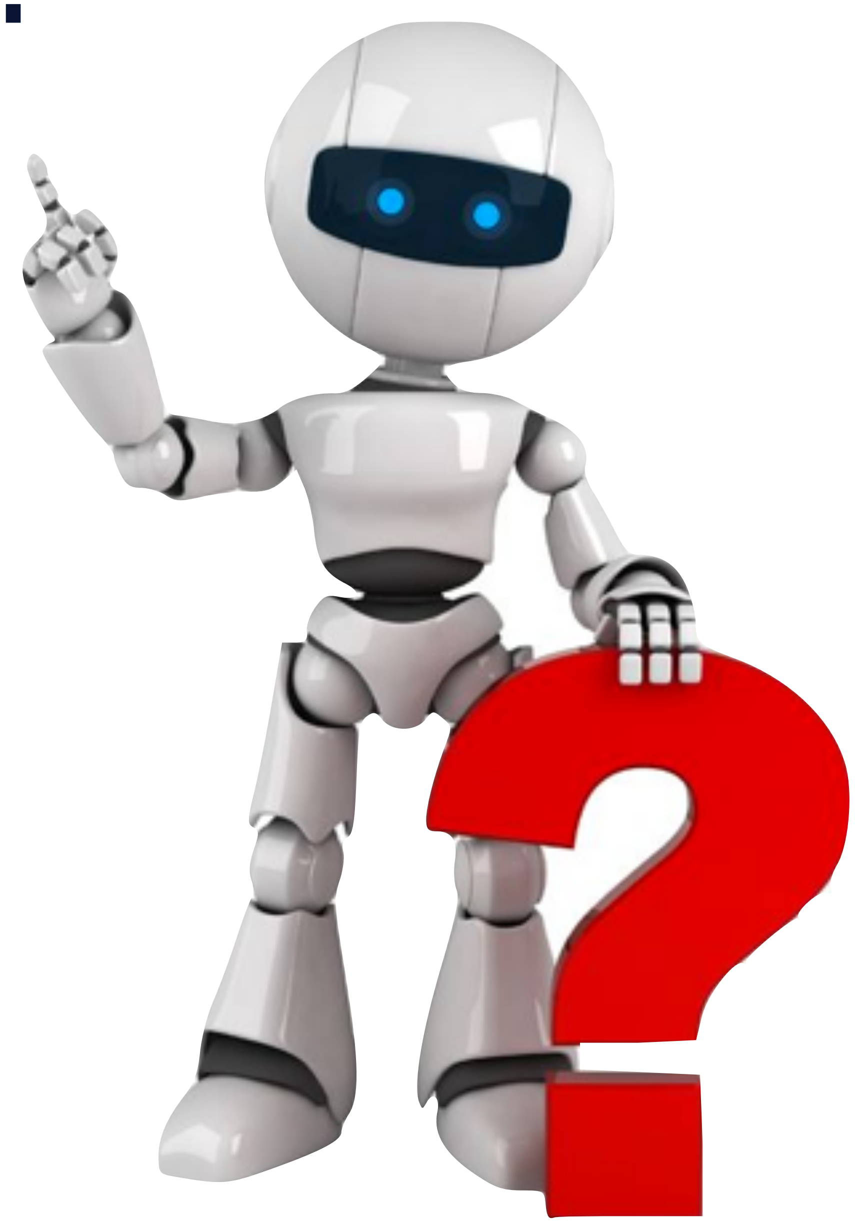


Image credit: http://www.globalrobots.com/

Institute for Aerospace Studies
UNIVERSITY OF TORONTO

STARS LABORATORY