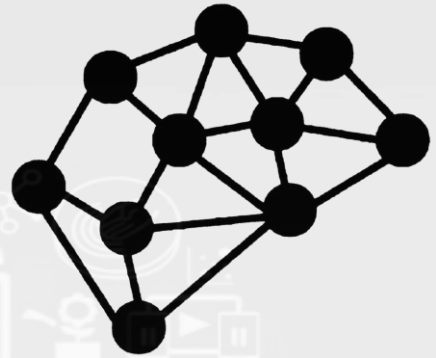
PURDUE UNIVERSITY | Elmore Family School of Electrical and Computer Engineering



Machine Intelligence & Networked Data Science

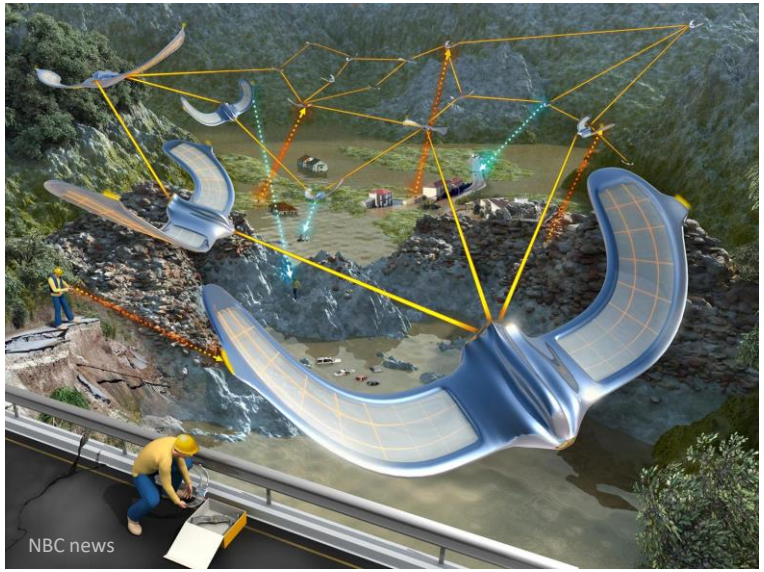
AI at Scale: Robustness and Security in Adversarial Environments

Abolfazl Hashemi

Assistant Professor of ECE

Machine Intelligence and Networked Data Science Lab **(MINDS)**

Collaborative Learning Systems



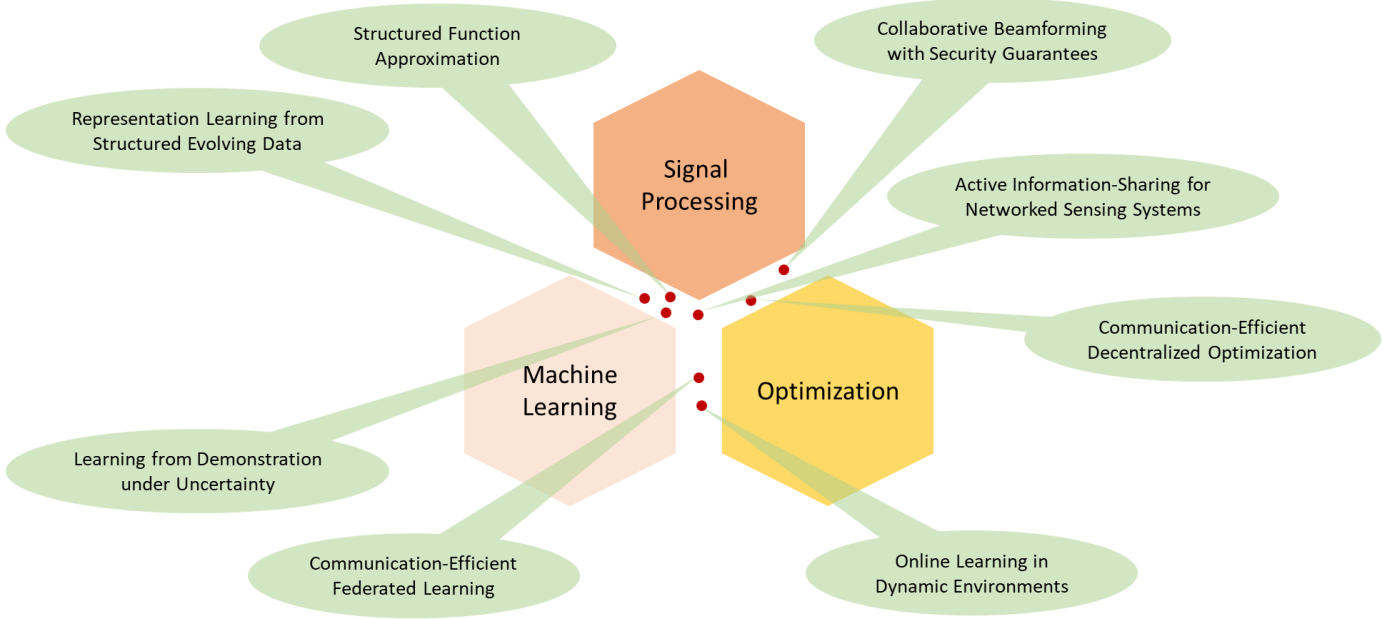
WIRED

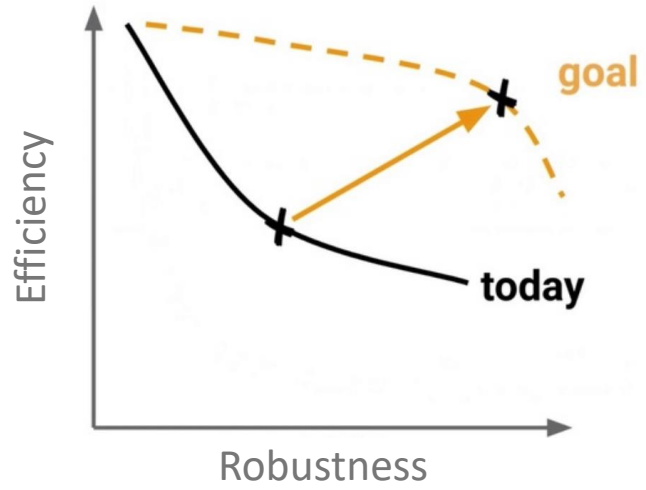
Hackers Remotely Kill a Jeep on the Highway—With Me in It

ANDY GREENBERG SECURITY 07.21.15 6:00 AM

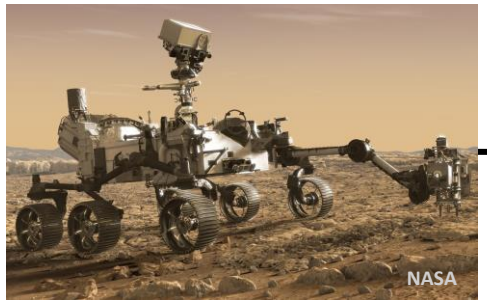
Update: Chrysler recalls 1.4M vehicles after Jeep hack

COMPUTERWORLD

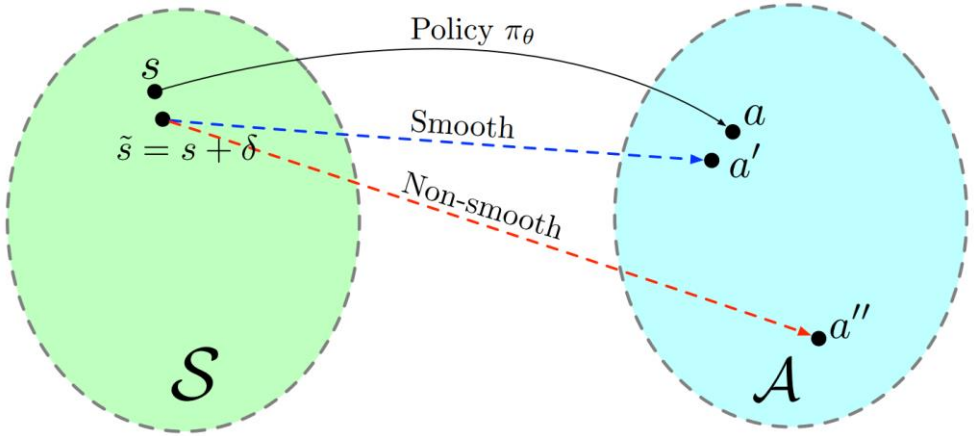


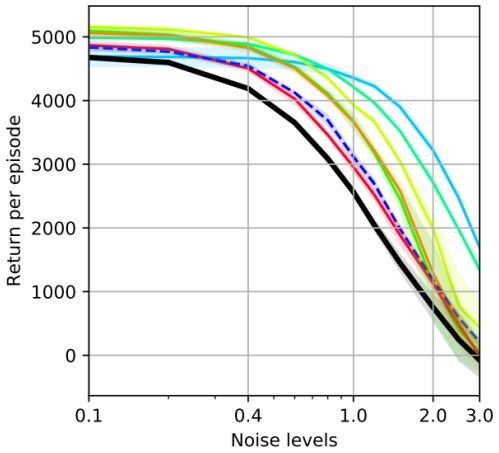


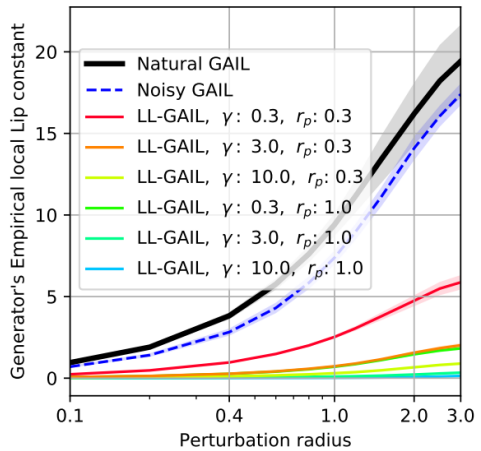
Online Learning in Adversarial Environment (AFOSR)

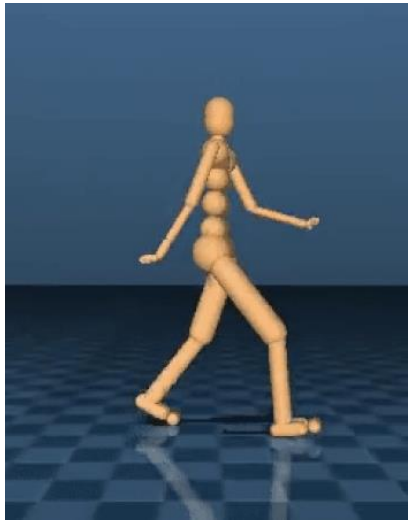


- Structured reward estimation and policy update
- Minimax optimal guarantees with high probability

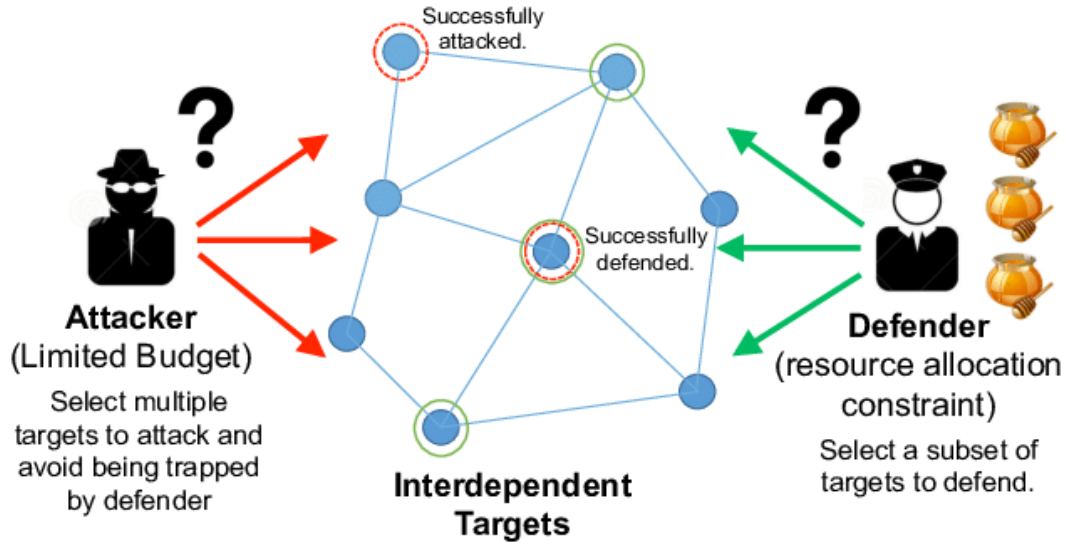




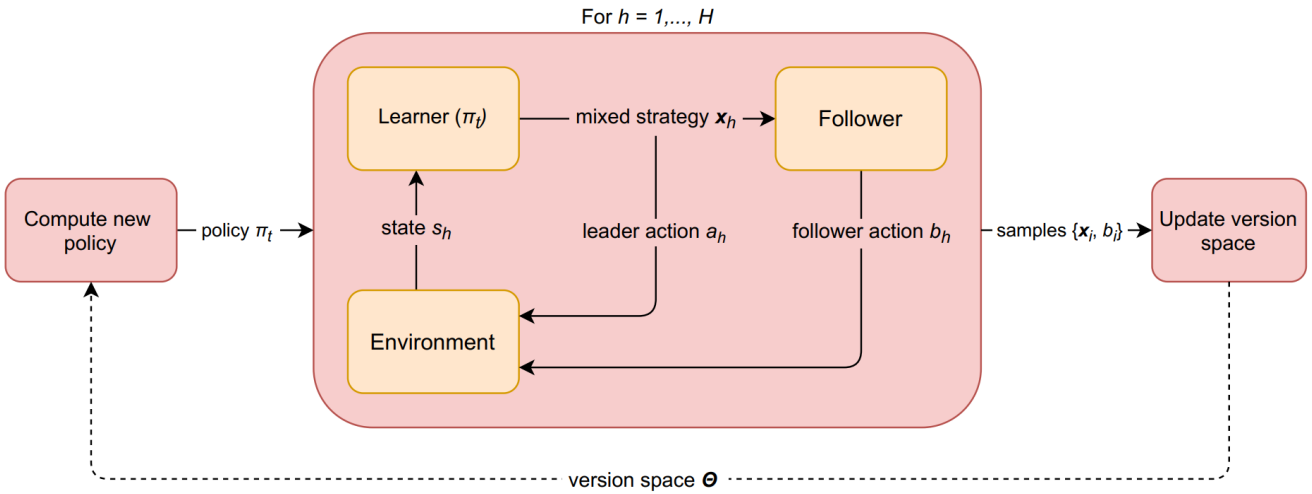


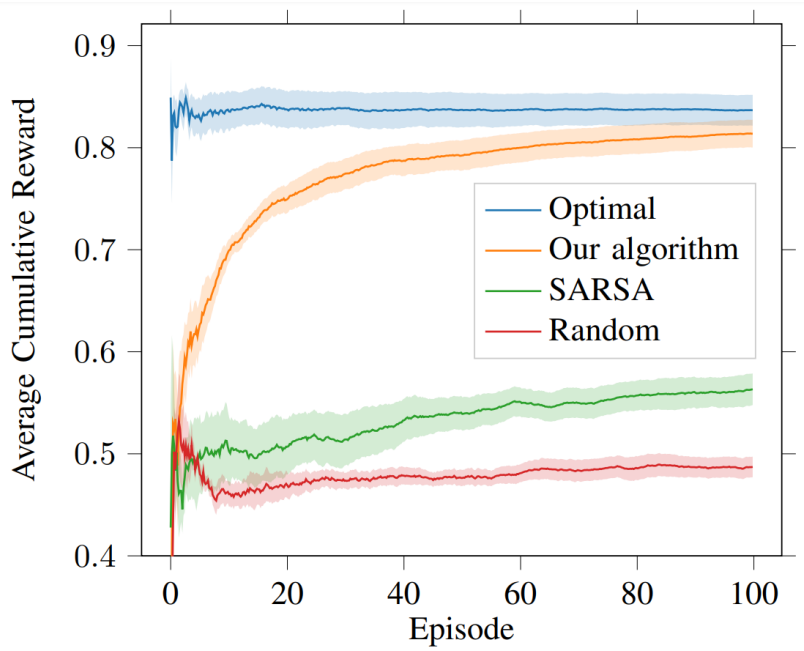


Dynamic Security Games for Resource Allocation

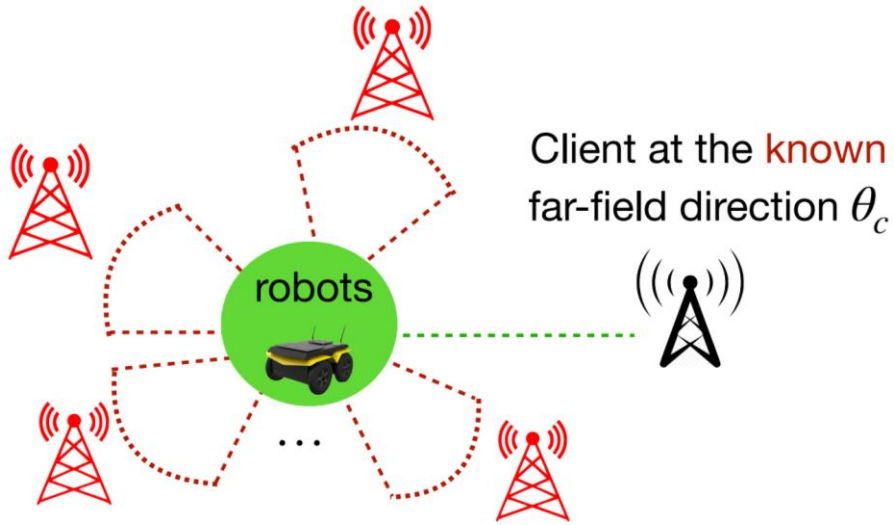


- Adaptive policy optimization
- Performance guarantees with high probability



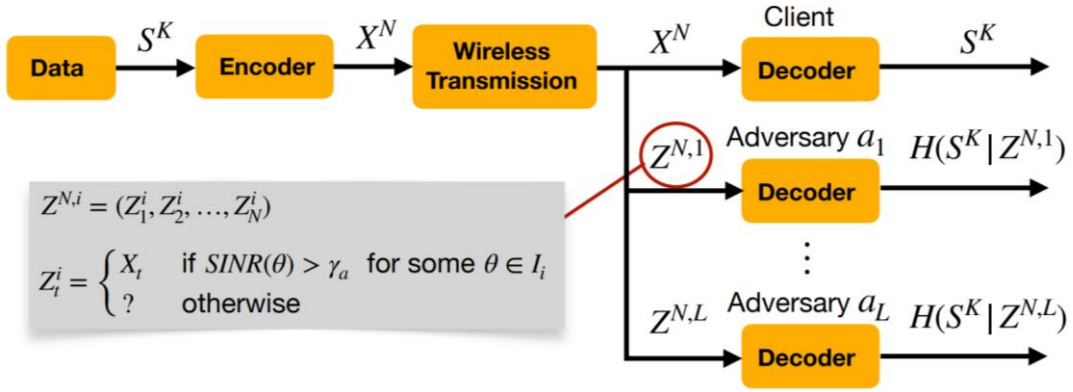


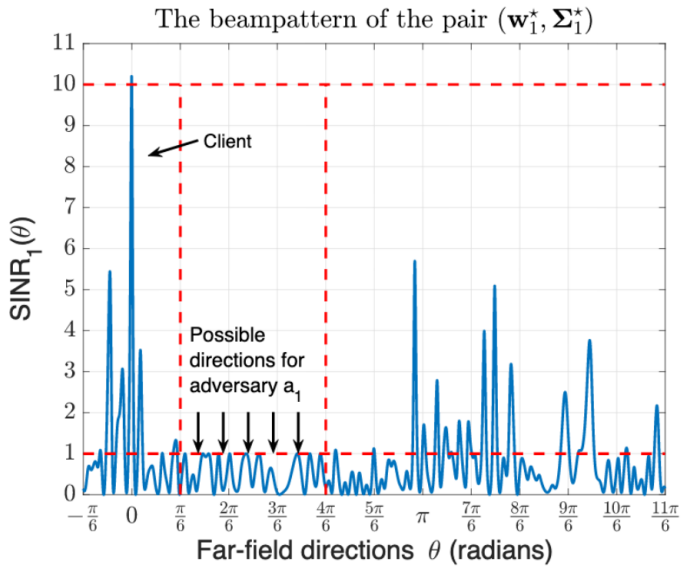
Physical-Layer Security via Distributed Beamforming (ARL)

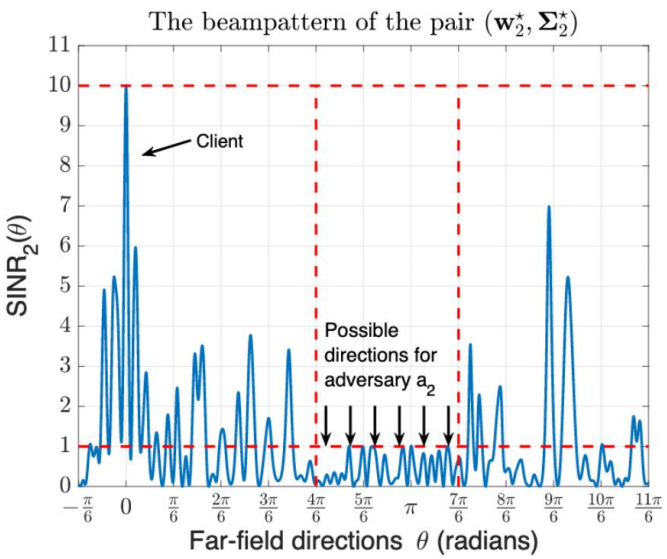


Semi-infinite non-convex program!

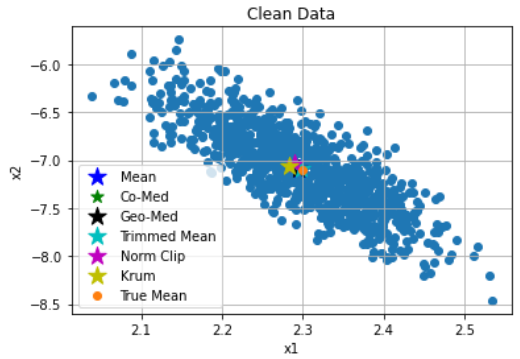
We present a **finite convex** program whose solution coincides with this program **with probabilistic guarantees**

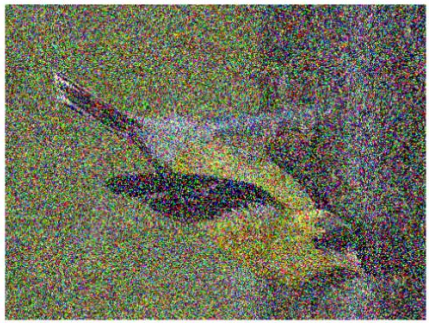


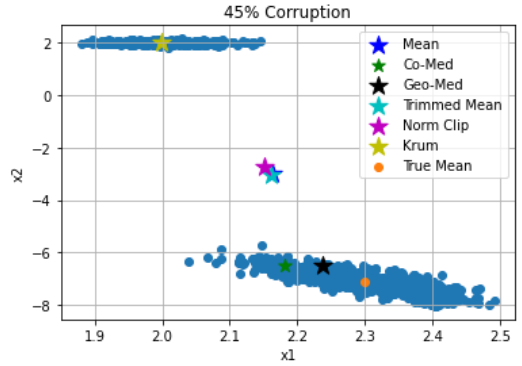


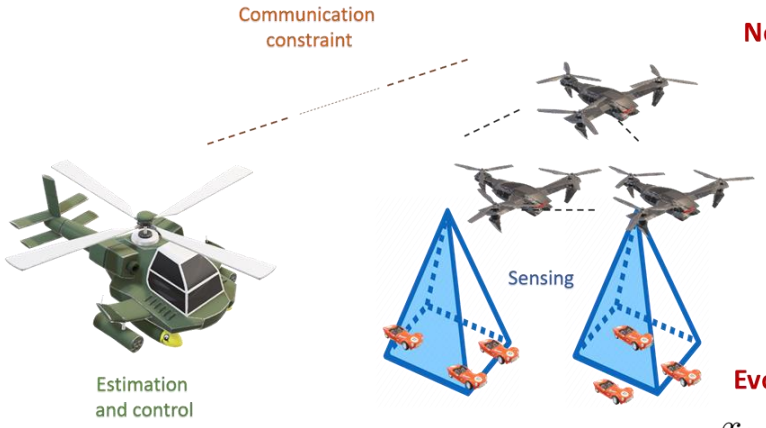


Robust Collaborative Learning in High Dimensions









Noisy and time-varying observation model

$$y_t = h(x_t) + \nu_t$$

Evolving unknown states

$$x_t = g(x_{t-1}, u_t) + v_t$$

- Exploiting low-dimensional structures
- Subspace sampling AND memory augmentation
- Performance guarantees for non-convex Tasks

	Corruption (%)	SGD	CMD	BGMD	GMD
ResNet18 - CIFAR10 (heterogeneous)					
Clean	-	82.29±1.32	85.50±1.43	84.82±0.76	**85.65**±0.48
Gradient Corruption					
Bit Flip	20	-	80.87±0.21	84.56±0.06	**88.07**±0.05
	40	-	77.41±1.04	**82.66**±0.31	80.81±0.01
Additive	20	20.7±1.56	54.75±0.38	**83.84**±0.12	82.40±0.90
	40	-	23.35±6.13	**82.79**±0.68	79.46±0.24