# Computationally Efficient Decision Making Under Uncertainty in High-Dimensional State Spaces

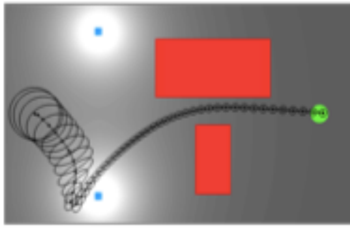
**Dmitry Kopitkov and Vadim Indelman**

TECHNION
Israel Institute of Technology

International Conference on Intelligent Robots and Systems (IROS), October 2016

# Introduction

- Decision making under uncertainty - fundamental problem in autonomous systems and artificial intelligence
- Examples
  - **Belief space planning** in uncertain/unknown environments (e.g. for autonomous navigation)
  - **Active simultaneous localization and mapping** (SLAM)
  - **Informative planning, active sensing**
  - **Sensor selection, sensor deployment**
  - **Multi-agent** informative planning and active SLAM
  - **Graph sparsification** for long-term autonomy



# Introduction

- Information-theoretic decision making
  - **Objective**: find action that minimizes an information-theoretic metric (e.g. entropy, information gain, mutual information)
  - **Problem types: Unfocused** (entropy of **all** variables), and **Focused** (entropy only of **subset** of variables)

- Decision making over **high-dimensional** state spaces is expensive!

$$X \in \mathbb{R}^n \qquad \Lambda \equiv \Sigma^{-1} \in \mathbb{R}^{n \times n}$$

- Evaluating action impact typically involves determinant calculation: $O(n^3)$ (smaller for sparse matrices)

- Existing approaches typically calculate posterior information (covariance) matrix for **each** candidate action, and then its determinant

# Key idea

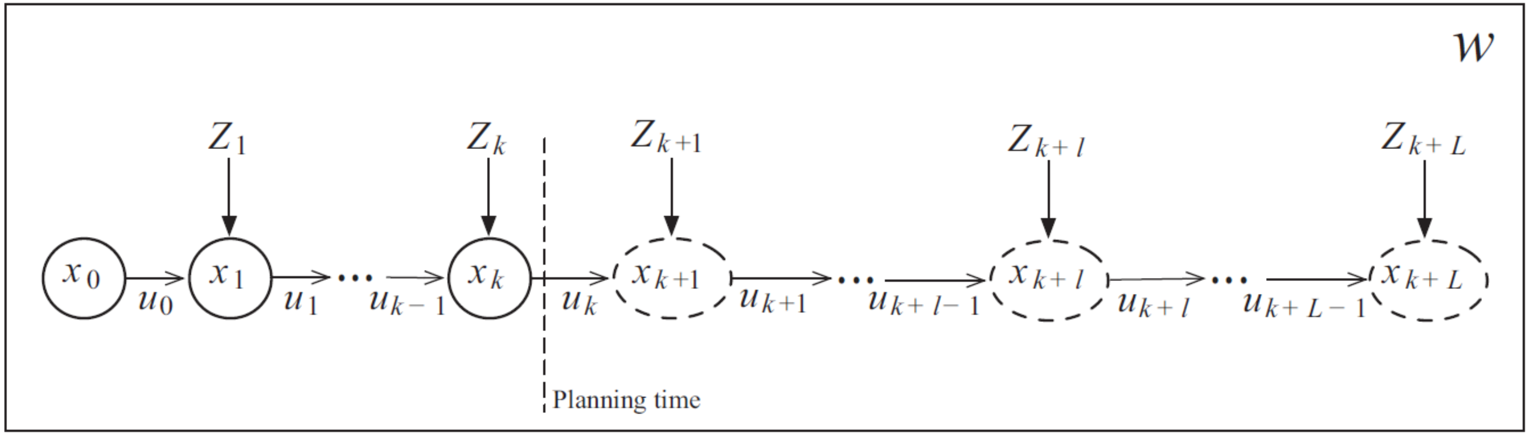
- Resort to **matrix determinant lemma and calculation re-use** techniques for **information-theoretic decision making**

$$\left|\Lambda_k + A^T A\right| = \left|\Lambda_k\right| \cdot \left|I_m + A \cdot \Sigma_k \cdot A^T\right| \qquad \Sigma_k \equiv \Lambda_k^{-1}$$

- **High-level overview:**
  - **Avoid** calculating determinants of large matrices
  - **Re-use** of calculations
  - Per-action evaluation **does not depend on state** dimension
  - Yet - **exact** and **general** solution

# Problem Formulation

- Given action $a = u_{k:k+L-1}$ and new observations $Z_{k+1:k+L}$ , future belief is:




$$b\left[X_{k+L}\right] = p\left(X_{k+L} | Z_{0:k+L}, u_{0:k+L-1}\right) \propto p\left(X_k | Z_{0:k}, u_{0:k-1}\right) \prod_{l=k+1}^{k+L} \underbrace{p\left(x_l | x_{l-1}, u_{l-1}\right)}_{\text{motion model}} \underbrace{p\left(Z_l | X_l^o\right)}_{\text{measurement likelihood}}$$

# Problem Formulation

$$b\left[X_{k+L}\right] = p\left(X_{k+L} | Z_{0:k+L}, u_{0:k+L-1}\right) \propto p\left(X_k | Z_{0:k}, u_{0:k-1}\right) \underline{\prod_{l=k+1}^{k+L} p\left(x_l | x_{l-1}, u_{l-1}\right) p\left(Z_l | X_l^o\right)}$$

- Information-theoretic cost (entropy): $J_{\mathcal{H}}\left(a\right) = \mathcal{H}\left(p\left(X_{k+L} | Z_{0:k+L}, u_{0:k+L-1}\right)\right)$
- Decision making: $a^\star = \arg\min_{a \in \mathcal{A}} J_{\mathcal{H}}\left(a\right)$ $\quad \mathcal{A} \doteq \left\{a_1, a_2, \ldots, a_N\right\}$
- Gaussian distributions: $\Lambda_{k+L} = \Lambda_k + A^T \underbrace{A}_{\text{Action Jacobian}}$

$$J_{\mathcal{H}(a)} = \frac{n}{2}\left(1 + \ln\left(2\pi\right)\right) - \frac{1}{2}\boxed{\ln\left|\Lambda_{k+L}\right|}$$

- Impact evaluation for a candidate action is in the general case: $O(n^3)$

# Information Gain (IG) & Matrix Determinant Lemma

- Use IG instead of entropy:

$$J_{IG}(a) \doteq \mathcal{H}(b[X_k]) - \mathcal{H}(b[X_{k+L}]) = \frac{1}{2} \ln \frac{\left|\Lambda_k + A^T A\right|}{\left|\Lambda_k\right|}$$

- **Applying matrix determinant lemma**:

$$\left|\Lambda_k + A^T A\right| = \left|\Lambda_k\right| \cdot \left|I_m + A \cdot \Sigma_k \cdot A^T\right|$$

$$\Sigma_k \equiv \Lambda_k^{-1} \in \mathbb{R}^{n \times n}$$

$$A \in \mathbb{R}^{m \times n}$$

$$\boxed{J_{IG}(a) = \frac{1}{2} \ln \left|I_m + A \cdot \Sigma_k \cdot A^T\right|}$$

**Typically:** $m \ll n$

**Examples: sensor deployment, active SLAM**

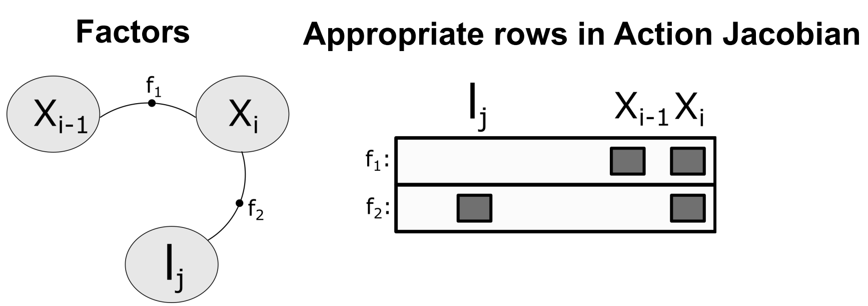
**Cheap, given $\Sigma_k$!**

# Sparse Structure of Action Jacobian

- Jacobian $A \in \mathbb{R}^{m \times n}$ of action $a$ represents new terms in the joint pdf:

$$b\left[X_{k+L}\right] = p\left(X_{k+L} | Z_{0:k+L}, u_{0:k+L-1}\right) \propto p\left(X_k | Z_{0:k}, u_{0:k-1}\right) \underline{\prod_{l=k+1}^{k+L} p\left(x_l | x_{l-1}, u_{l-1}\right) p\left(Z_l | X_l^o\right)}$$

- **Illustration** example:

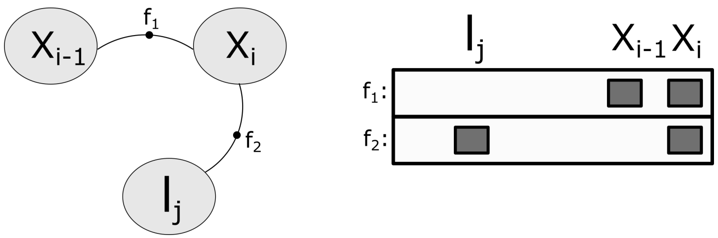



- Columns of $A$ that are not involved in new terms, contain only zeros

# Using Sparsity of Action Jacobian

$$J_{IG}(a) = \frac{1}{2}\ln\left|I_m + A \cdot \Sigma_k \cdot A^T\right| \quad \Longrightarrow \quad J_{IG}(a) = \frac{1}{2}\ln\left|I_m + A_C \cdot \Sigma_k^{M,C} \cdot A_C^T\right|$$

- $C$ is set of variables involved in $A$
- $A_C$ is constructed from $A$ by removing all zero columns
- $\Sigma_k^{M,C}$ is prior marginal covariance of $C$




- Only **few entries** from the prior covariance are actually required!

# Re-use of Calculation

- **Key observations**:
  - Given $\Sigma_k^{M,C}$, calculation of action impact depends only on action Jacobian:

$$J_{IG}(a) = \frac{1}{2} \ln \left| I_m + A_C \cdot \Sigma_k^{M,C} \cdot A_C^T \right|$$

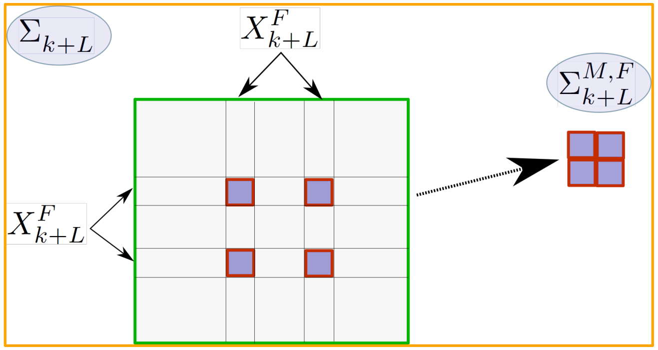
  - Different candidate actions often share many involved variables

- **Re-use calculations**:
  - Combine variables involved in all candidate actions into set $C_{All}$
  - Perform one-time calculation of $\Sigma_k^{M,C_{All}}$
    - Can be calculated efficiently from square root information matrix
    - Depends on state dimension $n$
  - Calculate $J_{IG}(a)$ for each action, using $\Sigma_k^{M,C_{All}}$

# Focused Decision Making

- Posterior entropy over focused variables $X^F_{k+L} \subseteq X_{k+L}$

$$J^F_{\mathcal{H}}(a) = \mathcal{H}\left(X^F_{k+L}\right) = \frac{n_F}{2}\left(1 + \ln(2\pi)\right) + \frac{1}{2}\ln\left|\Sigma^{M,F}_{k+L}\right|$$




- Applying **IG** and matrix determinant lemma of **Schur complement**:

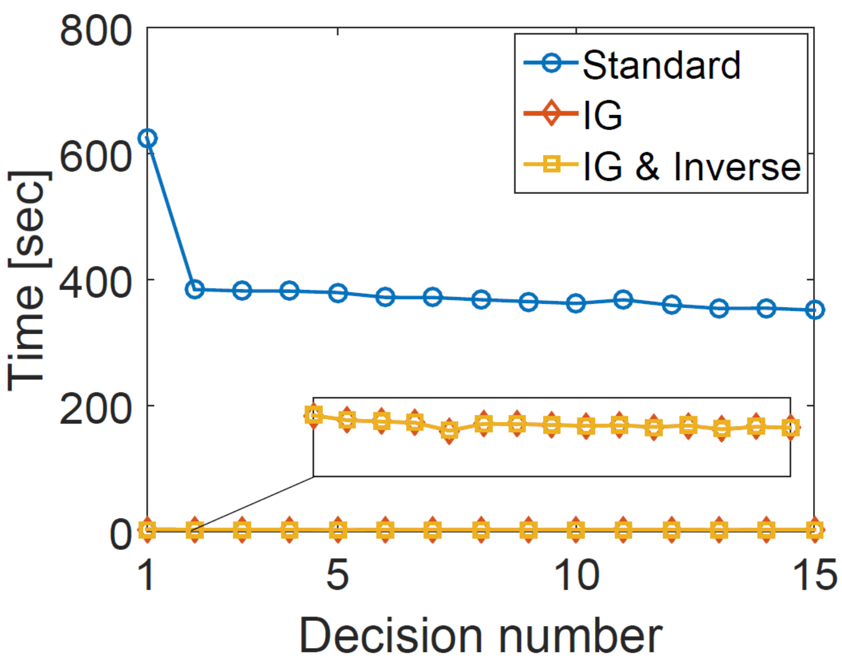
$$\boxed{\left|\Lambda_k\right| = \left|\Lambda^{M,F}_k\right| \cdot \left|\Lambda^R_k\right|} \Rightarrow \boxed{J^F_{IG}(a) = \frac{1}{2}\ln\frac{\left|I_m + A \cdot \Sigma_k \cdot A^T\right|}{\left|I_m + A_R \cdot \Sigma^{R|F}_k \cdot A^T_R\right|}}$$

$A_R$ – columns of $A$ related to $X^R_k$

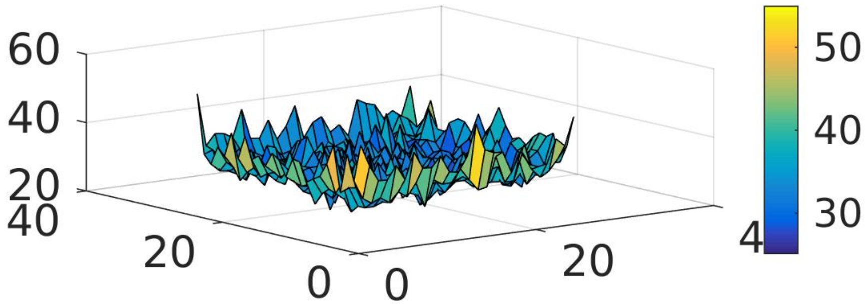
$\Sigma^{R|F}_k$ – is inverse of $\Lambda^R_k$

# Application to Sensor Deployment Problems

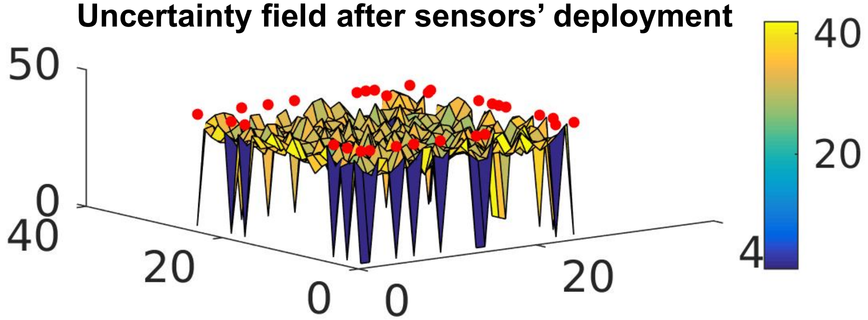
- Significant time reduction in *Unfocused* case




Uncertainty field (dense prior information matrix)




Uncertainty field after sensors' deployment

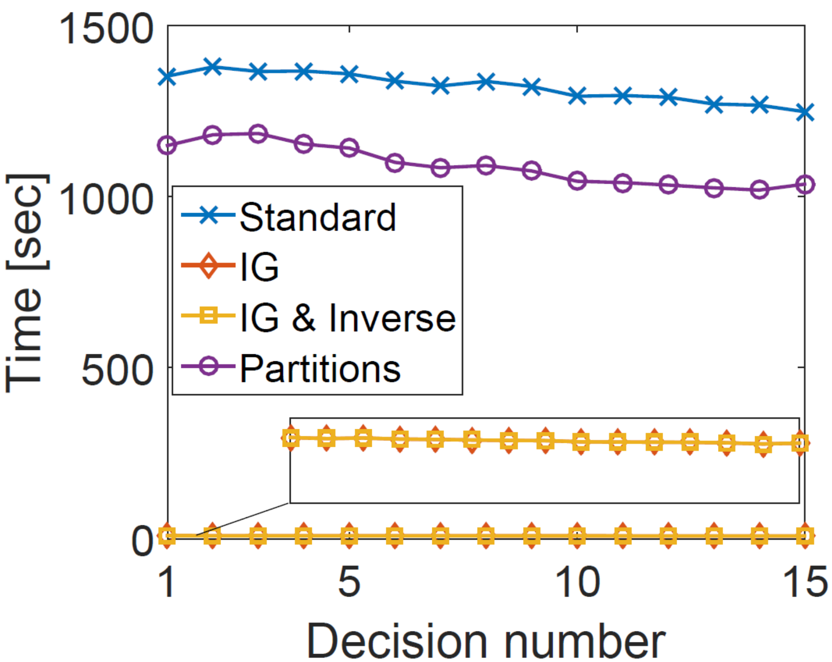


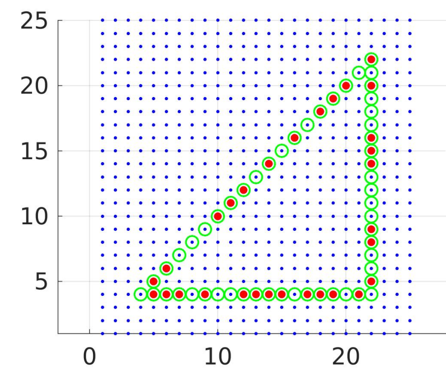



D. Kopitkov, V. Indelman, Computationally Efficient Decision Making Under Uncertainty in High-Dimensional State Spaces

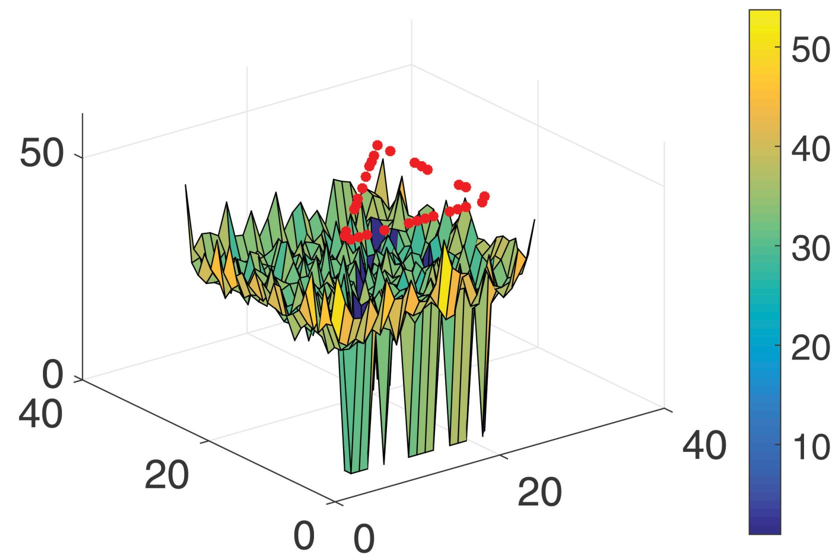
# Application to Sensor Deployment Problems

- Significant time reduction in *Focused* case

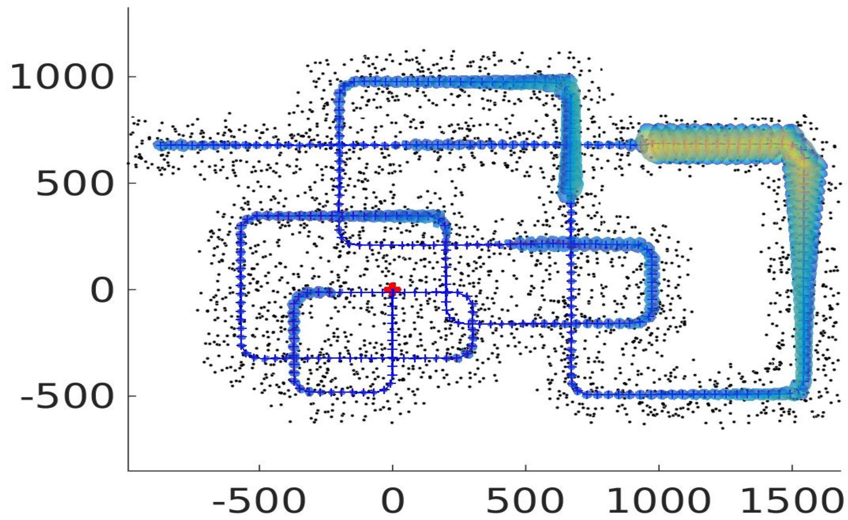


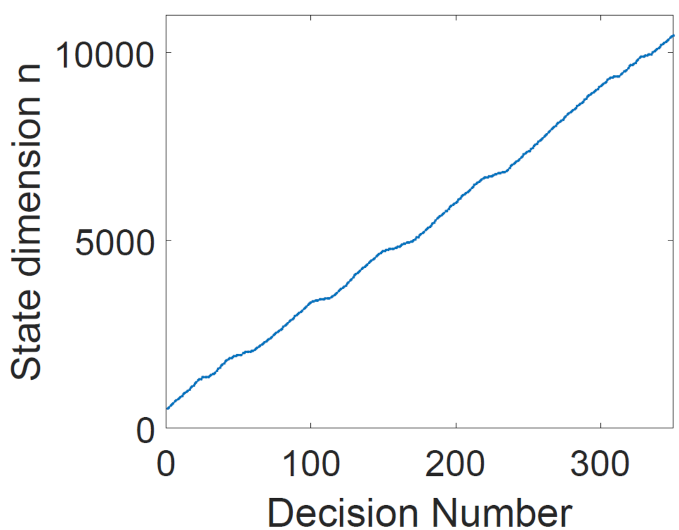

$$X^F_{k+L} \subseteq X_{k+L}$$

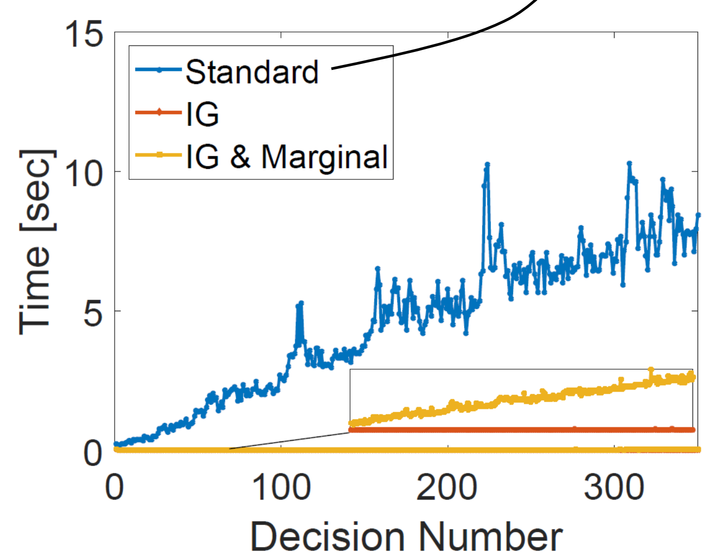


# Application to Measurement Selection (in SLAM Context)




**‘Standard’**: for each action, calculate posterior sqrt information matrix via **iSAM2**, then its determinant



# Conclusions

- **Decision making via matrix determinant lemma and calculation re-use**
  - Exact (no approximations applied)
  - General (any measurement model)
  - Per-candidate complexity does not depend on $n$
  - *Unfocused* and *Focused* problem formulations
  - Applicable to Sensor Deployment, Measurement Selection, Graph Sparsification, and many more..

- Multiple extensions to be investigated

TECHNION
Israel Institute of Technology