Connectivity and Localization in Wireless Sensor Networks

François Septier

Institut Mines-Télécom/Télécom Lille/LAGIS UMR CNRS 8219

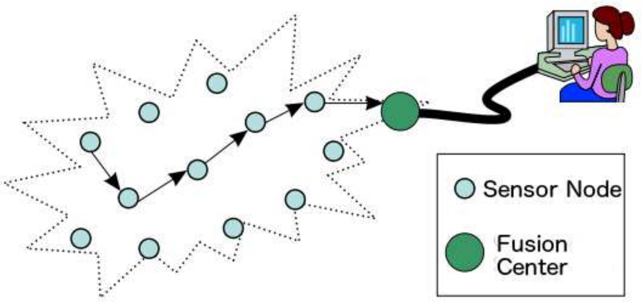
STM2014/CSM2014

July 28-31 2014, ISM

Wireless sensor network

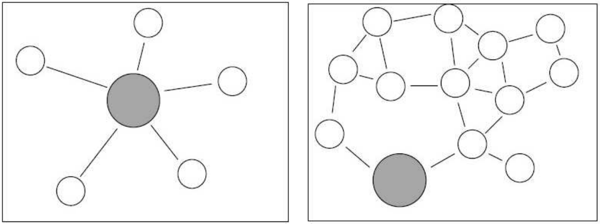
Spatially distributed autonomous sensors to

- monitor physical or environmental conditions (such as temperature, sound, pressure, etc.) ,
- cooperatively pass their data through the network to a main location.



WSN can be either :

- single-hop wireless transmission : *popular in short-range applications, such as smart homes*
- multi-hop wireless transmission (ad hoc) : *more interesting due to its high flexibility and ability to support long-range, large scale, and highly distributed applications*



After collecting information from the environment, sensors need to transmit aggregated data to gateways or Fusion Centers (FCs).

$\Rightarrow$ Important to ensure that every sensor can communicate with the FCs

Network connectivity

Outline

1. Connectivity in Wireless Sensor Network
   - Introduction to connectivity
   - System Model
   - Connectivity Study

2. Multiple Source Localization in WSN
   - Introduction
   - Problem Formulation
   - Proposed Bayesian solution

Outline

1. Connectivity in Wireless Sensor Network
   - Introduction to connectivity
   - System Model
   - Connectivity Study

2. Multiple Source Localization in WSN
   - Introduction
   - Problem Formulation
   - Proposed Bayesian solution

Introduction

A sufficient condition for reliable information transmission in WSNs is **full connectivity** of the network.

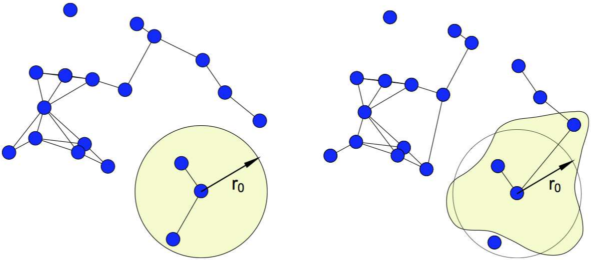
Definition Full connectivity

A network is said to be fully connected if every pair of nodes can communicate with each other, either directly or via intermediate relay nodes.

- Single-hop : Full connectivity is achieved if there exist a wireless link between each node and at least one gateways
- Multi-hop (ad hoc) : situation more complicated since each single node contributes to the connectivity of the entire network

  $\hookrightarrow$ *depends on spatial density, transmission/ reception capabilities, characteristic of the wireless channel, etc...*

1. Purely geometric link model : 2 nodes are connected if they are not further apart than a certain threshold distance r_0
   ⇝ (Chen et al., 1989), (Piret et al. , 1991), (Dousse et al. , 2006) [Percolation theory]
2. Shadow fading link model : consider the randomness nature of the wireless channel induced for example by shadowing effects that are caused by obstacles (more realistic)
   ⇝ (Orriss et al, 2003), (Bettstetter et al, 2004), (Dardari et al, 2007), (Zannella et al. , 2009)



a. Purely geometric link model b. Shadow fading link model

1. Purely geometric link model : 2 nodes are connected if they are not further apart than a certain threshold distance r_0
   ⇝ (Chen et al., 1989), (Piret et al. , 1991), (Dousse et al. , 2006) [Percolation theory]
2. Shadow fading link model : consider the randomness nature of the wireless channel induced for example by shadowing effects that are caused by obstacles (more realistic)
   ⇝ (Orriss et al, 2003), (Bettstetter et al, 2004), (Dardari et al, 2007), (Zannella et al. , 2009)

Propose some extensions of results from *Shadow fading link model* by incorporating random survival time of sensors due to power constraints and/or failure.

Joint work with Gareth Peters, Ido Nevat and Laurent Clavier

Spatial Node Distribution

Spatial distribution of the nodes is given by a homogeneous Poisson point process of density λ per unit area.

- Number of nodes N_Ω in space Ω of size $\|\Omega\|$ follows a Poisson distribution, i.e.

$$N_\Omega \sim \mathcal{P}o(\lambda\|\Omega\|)$$

- The random location of the n-th node is denoted by

$$\mathbf{x}_n | N_\Omega \sim \mathcal{U}[\Omega]$$

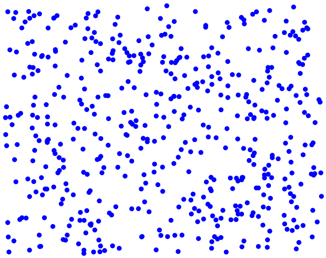


Illustration of Homogeneous Poisson distributed nodes

Wireless Channel Model

A wireless channel in which transmission of signal is subject to path-loss and shadowing is considered.
⇒ The power loss between the i-th and j-th nodes is a random variable (R.V.) defined as

$$L_{\mathbf{x}_i,\mathbf{x}_j} = k_0 + k_1 \log R(\mathbf{x}_i, \mathbf{x}_j) + S$$

with :

- k_0 and k_1 are known propagation constants
- $R(\mathbf{x}_i, \mathbf{x}_j)$ is the distance between the 2 randomly distributed nodes
- S is a R.V. representing the shadowing effect, which is generally assumed to be normal with zero mean and variance σ^2

Sensor Survival Time

Assumptions :

- The network is swithed on at time $t = 0$ and all nodes have a survival time (due to battery life and/or failure) relative to $t = 0$.
- The i-th node at location $\mathbf{x}_i$ has a survival time denoted by T_i which is considered as R.V. with distribution F_{T_i} - if $t < T_i$ then node i is active, otherwise this node is inactive.
- Survival times of each sensor are independent and identically distributed.

Connectivity Study

Definition 1 - l-audibility

At time $t = 0$, the intial l-audible set of nodes given the reference node location $\mathbf{x}_i \in \Omega$ is defined by

$$
\begin{aligned}
D_0(\mathbf{x}_i) &= \{\mathbf{x}_k : \mathbf{x}_k \in \Omega \text{ is } l \text{ audible with } \mathbf{x}_i\} \\
&= \{\mathbf{x}_k : \{L_{\mathbf{x}_i,\mathbf{x}_k} \leq l\} \cap \{\mathbf{x}_k \in \Omega\}, \forall k \in \{1, \ldots, N_\Omega\}\}.
\end{aligned}
$$

For $t > 0$ the number of audible nodes reduces due to the survival process that each node is subject to. This results in the l-audible set for a reference node at location $\mathbf{x}_i \in \Omega$ being defined at time t by

$$D_t(\mathbf{x}_i) = D_0(\mathbf{x}_i) \cap \{T_i > t\} \cap \{\mathbf{x}_k : T_k > t, \forall k \in \{1, \ldots, N_\Omega\}\}$$

The number of l-audible nodes at time t in some sub-region $A \subseteq \Omega$ is defined as :

$$N_A(t) = \sum_{k=1}^{N_A} \mathbb{1}\left[\mathbf{x}_k \in D_t(\mathbf{x}_i)\right]$$

Connectivity Study

Study of the probability density function of the distance between a pair of l-audible nodes at a given time t

$$f_{R_{\mathbf{x}_i,\mathbf{x}_j}(\tau)|\mathbf{x}_j \in D_0(\mathbf{x}_i)}(r) = \lambda(\tau)\delta_0(dr) + (1-\lambda(\tau))f_{R_{\mathbf{x}_i,\mathbf{x}_j}(t=0)|\mathbf{x}_j \in D_0(\mathbf{x}_i)}(r)$$

where

$$\lambda(\tau) = 1 - \overline{F}_{T_{\mathbf{x}_i}}(\tau)\overline{F}_{T_{\mathbf{x}_j}}(\tau) \quad \textit{[Complentary CDFs Survival Time]}$$

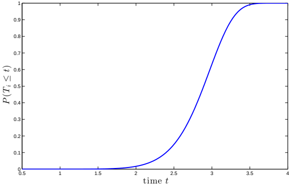
and

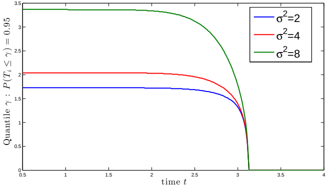
$$f_{R_{\mathbf{x}_i,\mathbf{x}_j}(t=0)|\mathbf{x}_j \in D_0(\mathbf{x}_i)}(r) = r \exp\left(-\frac{2}{k_1}\left(l - k_0 + \frac{\sigma^2}{k_1}\right)\right) \times \text{erfc}\left(\frac{k_0 - l + k_1 \log(r)}{\sqrt{2}\sigma}\right),$$

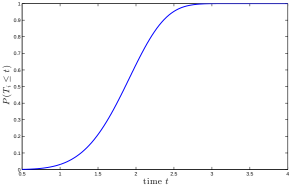
Study of the probability density function of the distance between a pair of l-audible nodes at a given time t [$l = 0,\ k_1 = 20,\ k_0 = -10$]

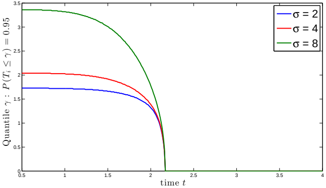
Survival Time $F_{T_i}(t)$

Quant. of Dist. betw. 2 connected nodes $\gamma(t) \mid \mathbb{P}(R(t) \leq \gamma(t)) = 0.95$









As the variance of the shadowing increases, node can establish links to neighbors that are further away.

Connectivity Study

Study of the probability distribution of the number of connected neighbors

- Number of l-audible neighbors in some region $A \subseteq \Omega$ of a reference node $\mathbf{x}_i$ is called its degree $N_A(\tau)$
- We derive its probability distribution, which is :

$$\begin{aligned} \mathbb{P}\left(N_A(\tau) = n\right) &= \sum_{n_A=n}^{\infty} \binom{n_A}{n} p(\tau)^n (1 - p(\tau))^{n_A - n} \mathbb{P}\left(N_A(0) = n_A\right) \\ &= \mathcal{P}o(p(\tau)\lambda\|A\|) \quad \textit{[From Thinning principle]} \end{aligned}$$

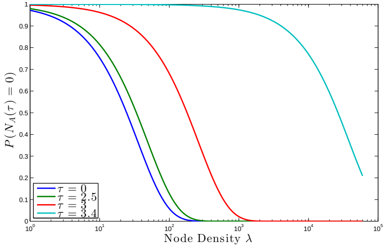
with

$$p(\tau) := \underbrace{\mathbb{P}\left(\mathbf{x}_k \in D_0(\mathbf{x}_i) | T_{\mathbf{x}_i} > \tau, T_{\mathbf{x}_k} > \tau\right)}_{\text{Proba of having a connected link}} \overline{F}_{T_{\mathbf{x}_k}}(\tau)\overline{F}_{T_{\mathbf{x}_i}}(\tau)$$

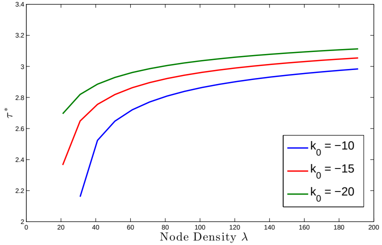
$\Rightarrow$ Closed-form expression derived

Study of the probability distribution of the number of connected neighbors
$\Rightarrow$ Useful quantity for network design !

Proba Isolated Node



$\tau^* = \inf\{\tau : \mathbb{P}(N_A(\tau) \leq 2) = 0.8\}$



- $\mathbb{P}(N_A(\tau) = 0) \nearrow$ when $\tau \nearrow$
- The minimum of time for which the number of connected neighbors becomes dramatically low (≤ 2) increases when the transmitted power or the node density increases

Perspectives

Conclusion

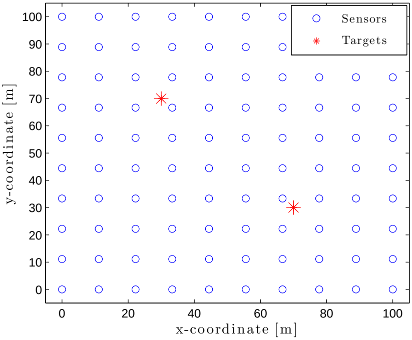
- Study the connectivity by taking into account the randomness of the wireless channel as well as the survival time of the sensor
  - Derivation of the pdf of the distance between two connected nodes
  - Derivation of the pdf of the number of nodes connected in some sub-region of the space.

On-going and Future works

- Derive some bounds on the probability of having a fully connected network by using the derived probability of having one isolated node ($\mathbb{P}(N_A(\tau) = 0)$)
- Consider that the sensor can recharge its battery
- Introduce some spatial dependency : Recharge can take longer in some region of the space (dark vs sunny regions)...

Outline

Wireless sensor networks (WSNs) : Composed of a large numbers of low-cost, low-power, densely distributed, and possibly heterogeneous sensors.



$\Rightarrow$ Makes possible energy emitting source !

Problem Formulation

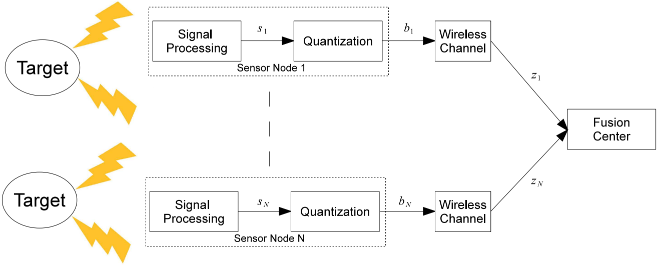
- Signal intensity measurements are very convenient and economical to localize a target,
  $\hookrightarrow$ no additional sensor functionalities and measurement features, such as direction of arrival (DOA) or time-delay of arrival (TDOA)
- Typical WSN has limited resources (energy and bandwidth) $\mapsto$ important to limit the communication with the network.
  $\hookrightarrow$ often desirable that only binary or multiple bit quantized data be transmitted from local sensors to the fusion center (processing node).
- The localization algorithm has to consider the imperfect nature of imperfect wireless channels between the local sensors and the fusion center

Existing Localization algorithm in WSN

1. Single source :
   - [Li & Hu, 2003] : a least -square method based on the energy ratios between sensors with analog measurements.
   - [Niu & Varshney, 2006] : a maximum likelihood using multi-bit sensor.
   - [Masazade et al., 2010] : importance sampler to approximate posterior distribution given quantized data.
   - [Ozdemir et al., 2010] : a maximum likelihood with imperfect communication channel and quantized data.
2. Multiple sources [known number of sources] :
   - [Sheng & Hu, 2005] :a maximum likelihood with perfect channel and analog measurement.

Aim : derive an inference algorithm for an unknown number of sources given quantized data with imperfect wireless channels (Param. of interest : K and $\mathbf{x}_K = \left[P_1, x_1, y_1, \ldots, P_K, x_K, y_K\right]^T$).

Joint work with Gareth Peters, Ido Nevat and T. L. Thu Nguyen



Received Signal at the i-th sensor

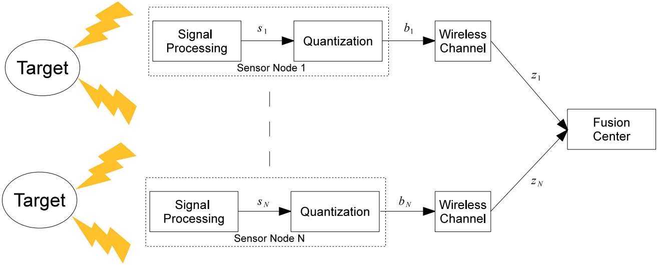
$$s_i = a_i + n_i$$

where the measurement noise, n_i, is Gaussian noise, i.e., $n_i \sim \mathcal{N}(0, \sigma^2)$ and

$$a_i = \sum_{k=1}^{K} P_k^{1/2} \left(\frac{d_0}{d_{i,k}} \right)^{\frac{n}{2}}$$

P_k : k-th source signal power at a reference distance d_0

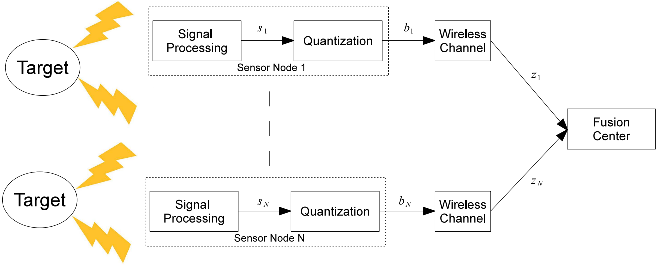
n Signal decay and $d_{i,k}$ distance between the i-th sensor and the k-th target



Quantization Stage at the i-th sensor
Transforms its input s_i to its output b_i through a mapping : $\mathbb{R} \mapsto \{0, \dots, L-1\}$ such that

$$b_i = \begin{cases} 0 & \lambda_{i,0} \leq s_i < \lambda_{i,1} \\ 1 & \lambda_{i,1} \leq s_i < \lambda_{i,2} \\ \vdots & \vdots \\ L-1 & \lambda_{i,L-1} \leq s_i < \lambda_{i,L} \end{cases}$$

with $\lambda_{i,0} = -\infty$ and $\lambda_{i,L} = +\infty$.



Wireless Communication from the i-th sensor to the FC

Quantized observation is transmitted to the fusion center through an imperfect channel which may introduce transmission errors.

The probability of a received observation z_i taking a specific value j, given the targets' parameters, $\mathbf{x}$, can be written as :

$$p(z_i = j|\mathbf{x}) = \sum_{m=0}^{L-1} \underbrace{p(z_i = j|b_i = m)}_{\text{known channel statistics}} p(b_i = m|\mathbf{x}) \quad (1)$$

Bayesian Framework

In this work, we are interested in estimating :

- unknown number of sources in the region, K^*
- the K^* sources' parameters (locations and transmitted powers)

$\Leftrightarrow$ joint model selection and parameter estimation problem

Indeed, we have :

- a collection of K competing models $\{\mathcal{M}_k\}_{k\in\{1,\ldots,K\}}$ (which corresponds to the number of sources)
- a vector of parameters associated with each model $\mathbf{x}_k = \left[P_1, x_1, y_1, \ldots, P_k, x_k, y_k\right]^T$

$\Rightarrow$ Propose a Bayesian solution

Bayesian Framework

Bayesian procedure proceeds from :

- a prior distribution over the collection of models, $p(\mathcal{M}_k)$,
- a prior distribution for the parameters of each model, $p(\mathbf{x}_k|\mathcal{M}_k)$,
- a likelihood distribution $p(\boldsymbol{z}|\mathbf{x}_k, \mathcal{M}_k)$

Thus,

1. **Model choice** one typically employs the maximum a posteriori (MAP)

$$k^* = \arg\max_k \{p(\mathcal{M}_k|\boldsymbol{z})\} = \arg\max_k \{p(\boldsymbol{z}|\mathcal{M}_k)p(\mathcal{M}_k)\}$$

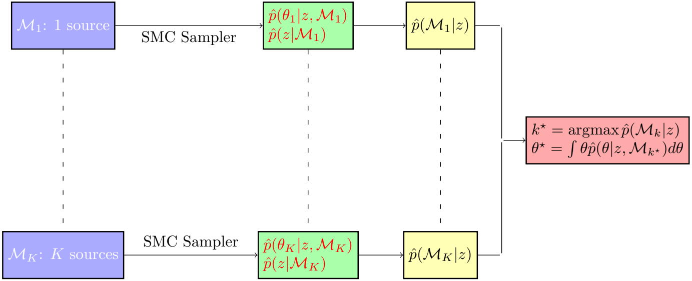
   where

$$p(\boldsymbol{z}|\mathcal{M}_k) = \int_{\Theta_k} p(\boldsymbol{z}|\mathbf{x}_k, \mathcal{M}_k)p(\mathbf{x}_k|\mathcal{M}_k)d\mathbf{x}_k$$

2. **Param. Estimate** The estimate of the parameters can be deduced from the posterior distribution associated with the model $\mathcal{M}_{k^*}$, i.e. $p(\mathbf{x}_{k^*}|\boldsymbol{z}, \mathcal{M}_{k^*})$

Unfortunately both $p(\boldsymbol{z}|\mathcal{M}_k)$ and $p(\mathbf{x}_{k^*}|\boldsymbol{z}, \mathcal{M}_{k^*})$ are intractable !

$\Rightarrow$ Propose to use advanced Monte-Carlo methods (SMC sampler) in order to have an accurate approximation of both quantities.



Derive an Sequential Monte Carlo sampler algorithm :

- Sequential algorithm which are able to deal with complex high-dimensional and/or multimodal posterior distribution
  - by using MCMC methodology
  - by introducing a sequence of progressive annealed distribution

(start with a distribution easy to sample from to the posterior of interest)

- produces a set of weighted samples that approximates the posterior distribution $p(\mathbf{x}_k|\boldsymbol{z},\mathcal{M}_k)$ and gives an unbiased estimate of $p(\boldsymbol{z}|\mathcal{M}_k)$
- *more details about this algorithm in my tomorrow's talk*

Single Source Scenario

Other parameters of the scenario :

- a signal decay exponent $n = 2$,
- a reference distance as and $d_0 = 1$,
- the region of interest is $100 \times 100m$ field
- the sensors are uniformly deployed in a grid .

Single Source Scenario

		SMC Recycling	Importance Sampler [Masazade et al., 2010]
	$N = 50$	0.0647 (0.0160)	0.1563 (0.1026)
25	$N = 100$	0.0527 (0.0112)	0.1181 (0.0870)
Iter.	$N = 200$	0.0456 (0.0082)	0.0943 (0.0715)
	$N = 50$	0.0543 (0.0131)	0.1159 (0.0796)
50	$N = 100$	0.0449 (0.0084)	0.0908 (0.0541)
Iter.	$N = 200$	0.0399 (0.0064)	0.0737 (0.0601)
	$N = 50$	0.0456 (0.0077)	0.0900 (0.0589)
100	$N = 100$	0.0406 (0.0073)	0.0735 (0.0413)
Iter.	$N = 200$	0.0367 (0.0053)	0.0611 (0.0427)

Accuracy to approximate the posterior distribution $p(x_1|\boldsymbol{z})$ in terms of the Kolmogorov-Smirnov distance (mean and standard deviation in parentheses).

$\Rightarrow$ **Significant improvement compared to existing IS algo. !**

Multiple Source Scenario

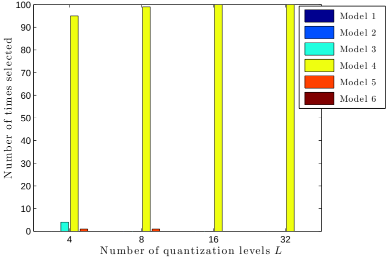
Parameters of the scenario :

- 4 sources in the ROI
- a signal decay exponent $n = 2$,
- a reference distance as and $d_0 = 1$,
- the region of interest is $100 \times 100m$ field
- the sensors are uniformly deployed in a grid .

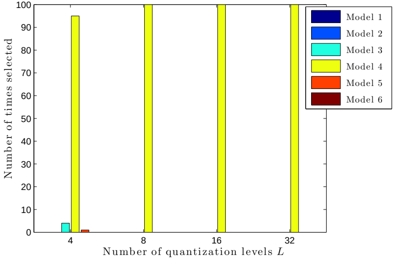
Multiple Source Scenario

Accuracy on the model choice :

$\sigma^2 = 1$



$\sigma^2 = 0.05$



Number of times that each model has been selected with the approximated model posterior from the SMC sampler with different number of quantization levels

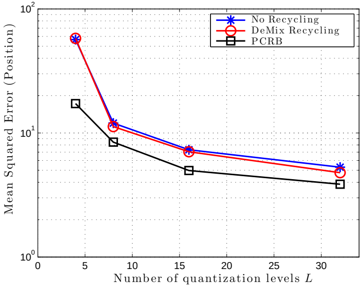
$\Rightarrow$ **Proposed algorithm clearly able to detect that there are 4 sources in the ROI !**

Numerical Simulations

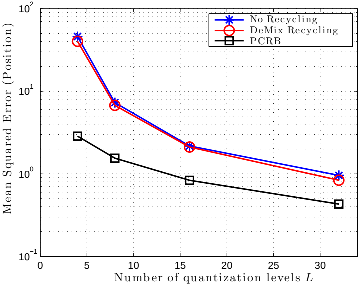
Multiple Source Scenario

Accuracy on the source localization :

$\sigma^2 = 1$



$\sigma^2 = 0.05$



MSE for the source locations with $\neq$ number of quantization levels L

We derive the posterior Cramér-Rao bound for this problem

- $\Rightarrow$ As expected, the accuracy on the localization improves as the number of quantization levels $\nearrow$
- $\Rightarrow$ Empirically demonstrate the good localization performance of the proposed algorithm

Conclusion and Future Works

Conclusion

- Propose efficient Bayesian algorithm to
  - estimate the number of source in the region of interet
  - estimate their locations as well as their transmitted powers
- Derive the posterior Cramér-Rao bound associated to the sources' parameters estimation

Future works

- Optimal sequential sensor selection scheme to avoid the transmission of information from all the sensors
- Utilize the derived posterior Cramér-Rao bound to optimize
  - placement of the sensors
  - quantization thresholds of the sensors