# A DYNAMIC BAYESIAN NETWORK APPROACH FOR DEVICE-FREE RADIO VISION: MODELING, LEARNING AND INFERENCE FOR BODY MOTION RECOGNITION

Stefano Savazzi, Sanaz Kianoush, Vittorio Rampa

Institute of Electronics, Computer
and Telecommunication Engineering,
Wireless Communication Systems Group (WCS)

Consiglio Nazionale delle Ricerche (C.N.R.)

# Outline

- Introduction to Radio Vision technologies for non-cooperative device-free localization and motion recognition

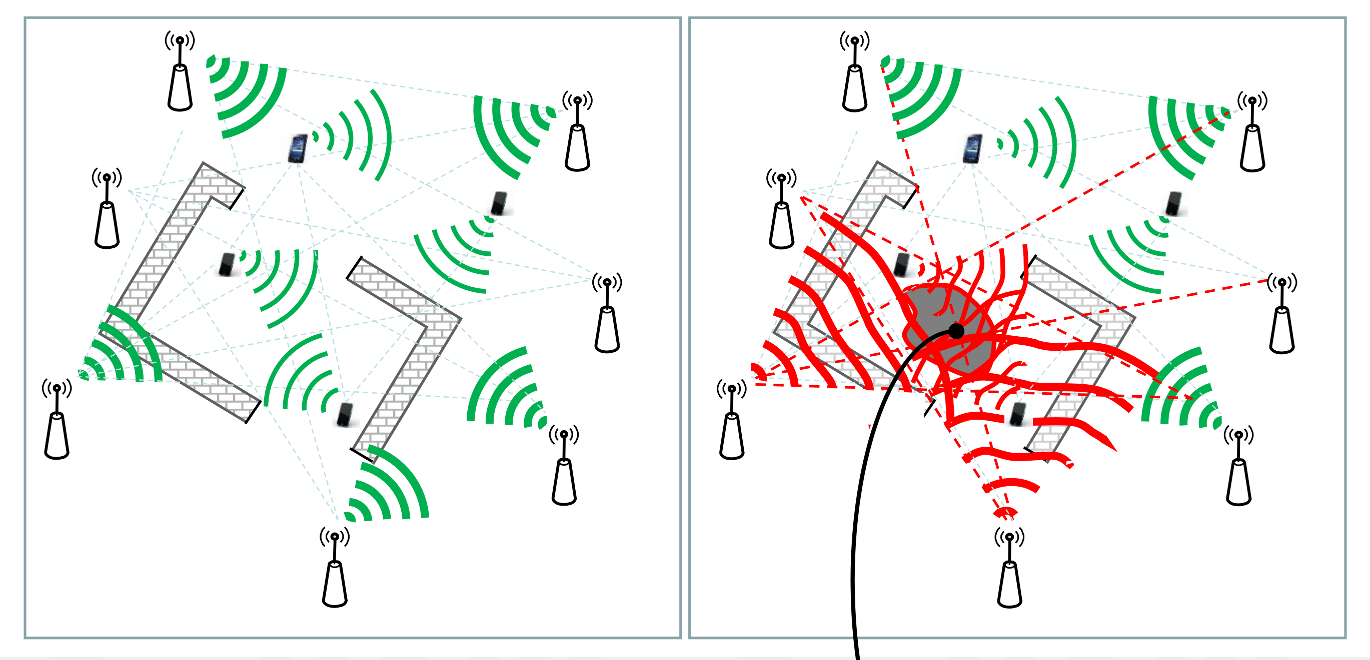
- Time-varying dynamic Bayesian network model for non-cooperative and device-free body motion recognition
  - Coupled hidden Markov (CHM) chain modeling of RF perturbations over co-located link pairs
  - Time-varying Bayesian network model for tracking coupled/uncoupled links

- Learning and classification problems discussed based on experimental measurements in a representative indoor environment
  - Applications: arm motion recognition, fall detection

# Device-free Radio Vision

- Augmented functionality provided by densely networked radio devices that monitor the fluctuations of a modulated RF field (e.g., adopted for wireless communications)

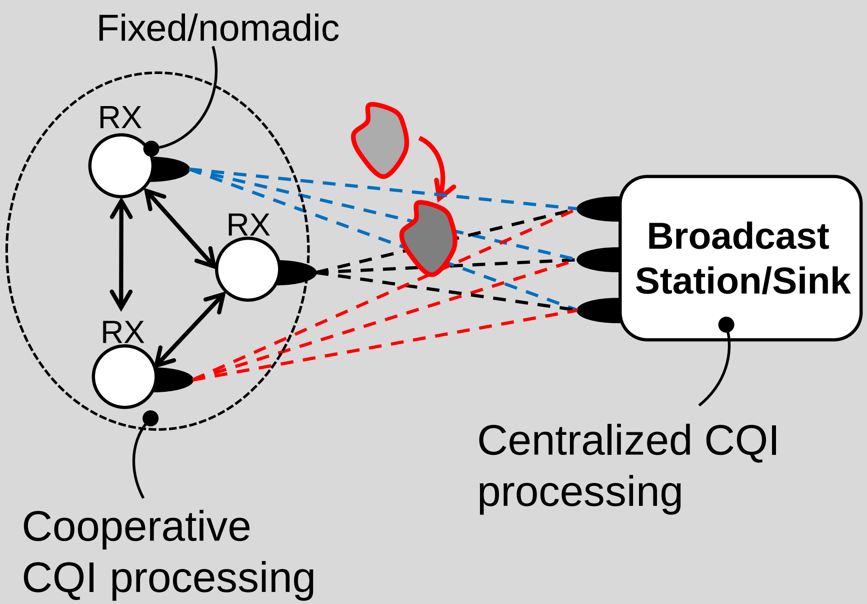
- Radio-vision leverage body-induced diffraction, reflection and scattering phenomena that affect RF propagation for ubiquitous human-scale sensing



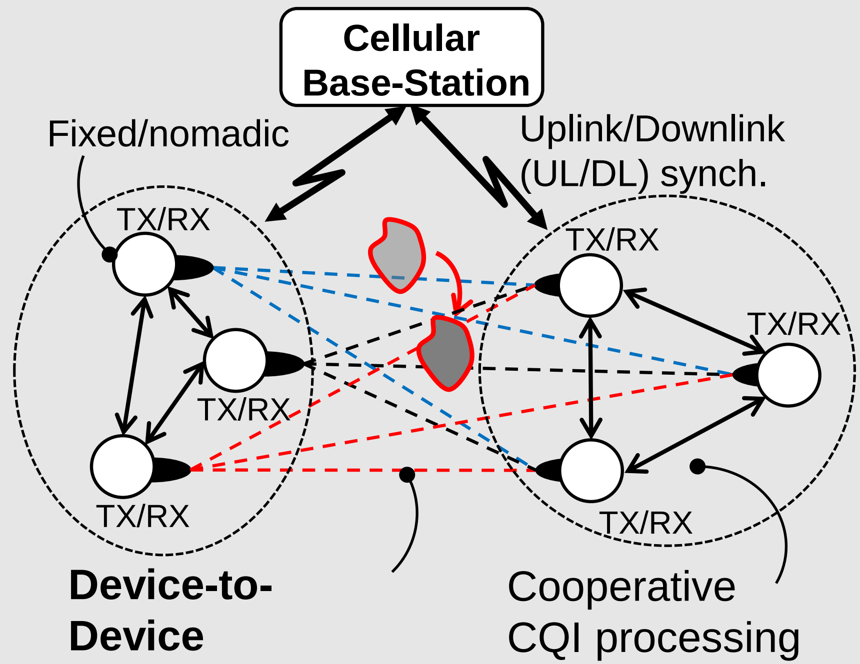
**Human-induced fading** (diffraction/reflection/ scattering)

# Radio Vision: typical configurations

**Passive** – Capturing Ambient Radio Signals

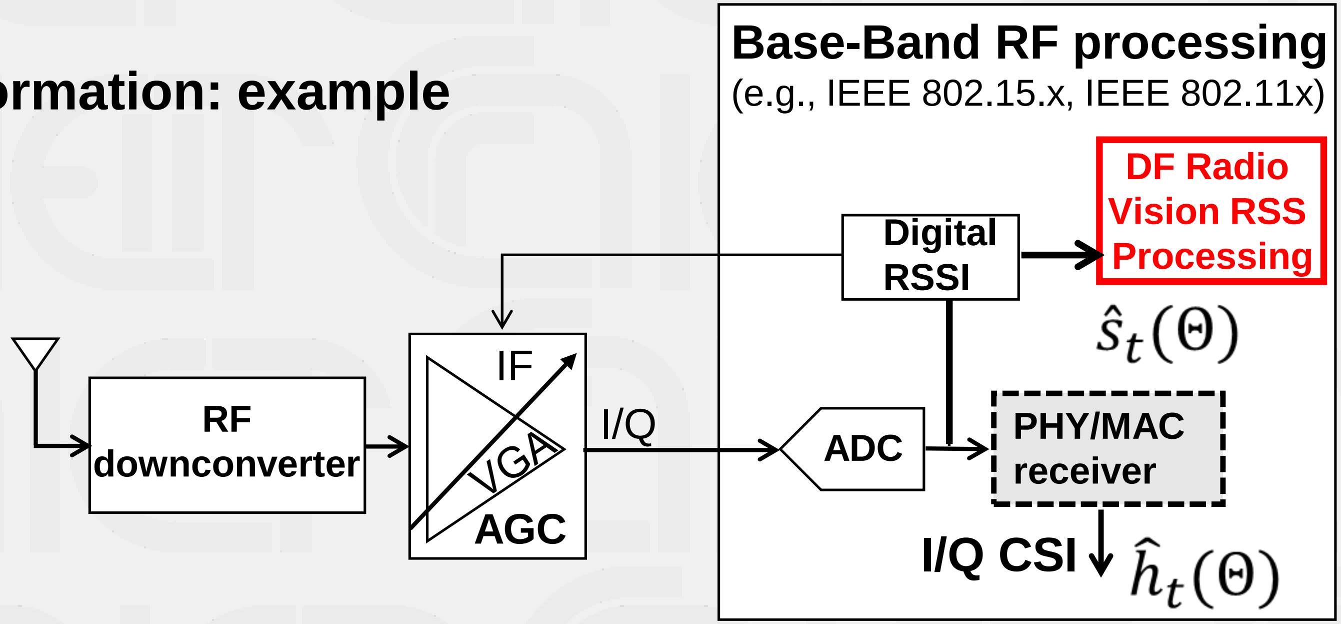



**Active** – Device-to-Device Communications

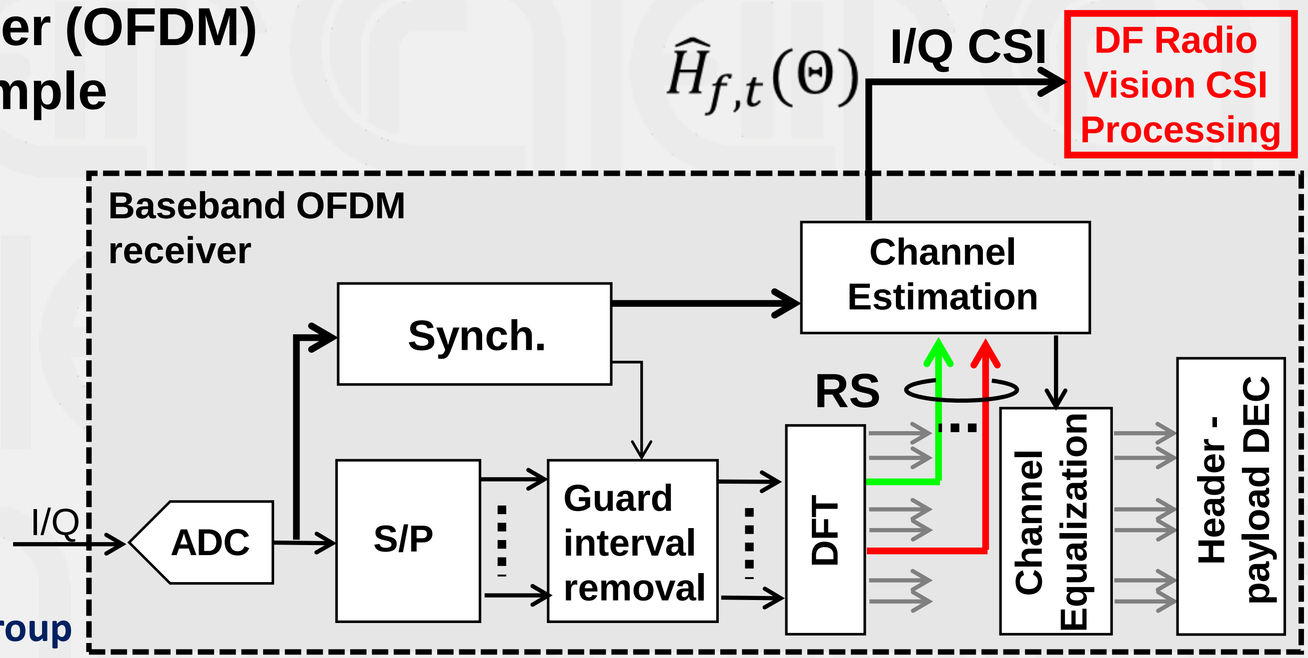


# Channel Quality Information extraction

(a) **RSS information: example**



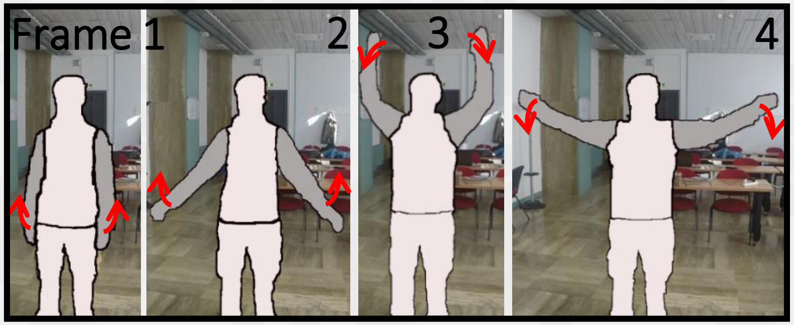

(b) **CSI in multi-carrier (OFDM) modulation: example**



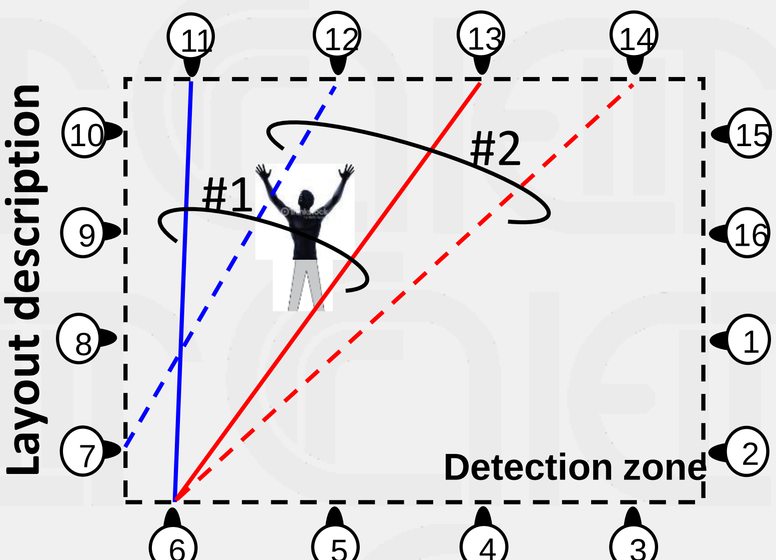
# Device-free Activity Recognition:

## Gesture and motion detection

**Gesture detection (case study, example)**




- Wireless mesh network whose transmissions are organized into periodic frames, or symbols
- A person inside the propagation area performing an activity, in a pre-assigned location, defined as a generic combinations of elementary body motions (arm motions in the example)




- **Time-Varying Dynamic Bayesian Network (TV-DBN)** introduced to describe the joint human-induced RF fluctuations among co-located links and account for spatial (link-wise) correlation of the channel response over multiple links:

1. “link hidden states”, CQI shifts and profile definition
2. prior Bayes network structure (defines initial dependency of link states)
3. Transition network graph
4. Time-varying transition network graph (to model time-varying coupled and uncoupled link states)

# Link hidden states, CQI shifts and profile

The effects of the user state $\Theta$ on the channel response are observed over $\mathrm{T}$ consecutive received frames from which CQI information can be extracted.

The CQI (in dB scale) over links $\ell_1, \ldots, \ell_N$ and frame (symbol) $1 \leq t \leq \mathrm{T}$ corresponding to the user state $\Theta$ can be modeled as

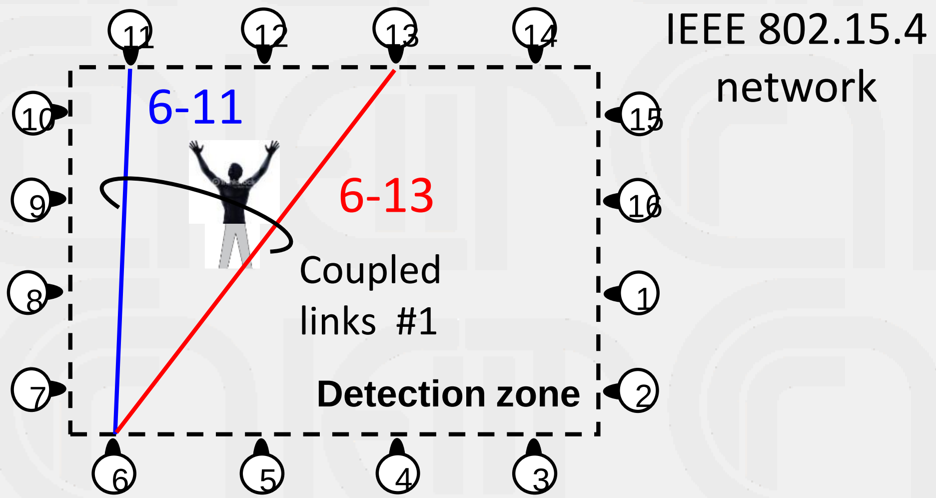
$$\mathbf{s}_t(\Theta) = \mathbf{q}_t(\Theta) + \bar{\mathbf{s}}(\varnothing) + \mathbf{u}_t$$

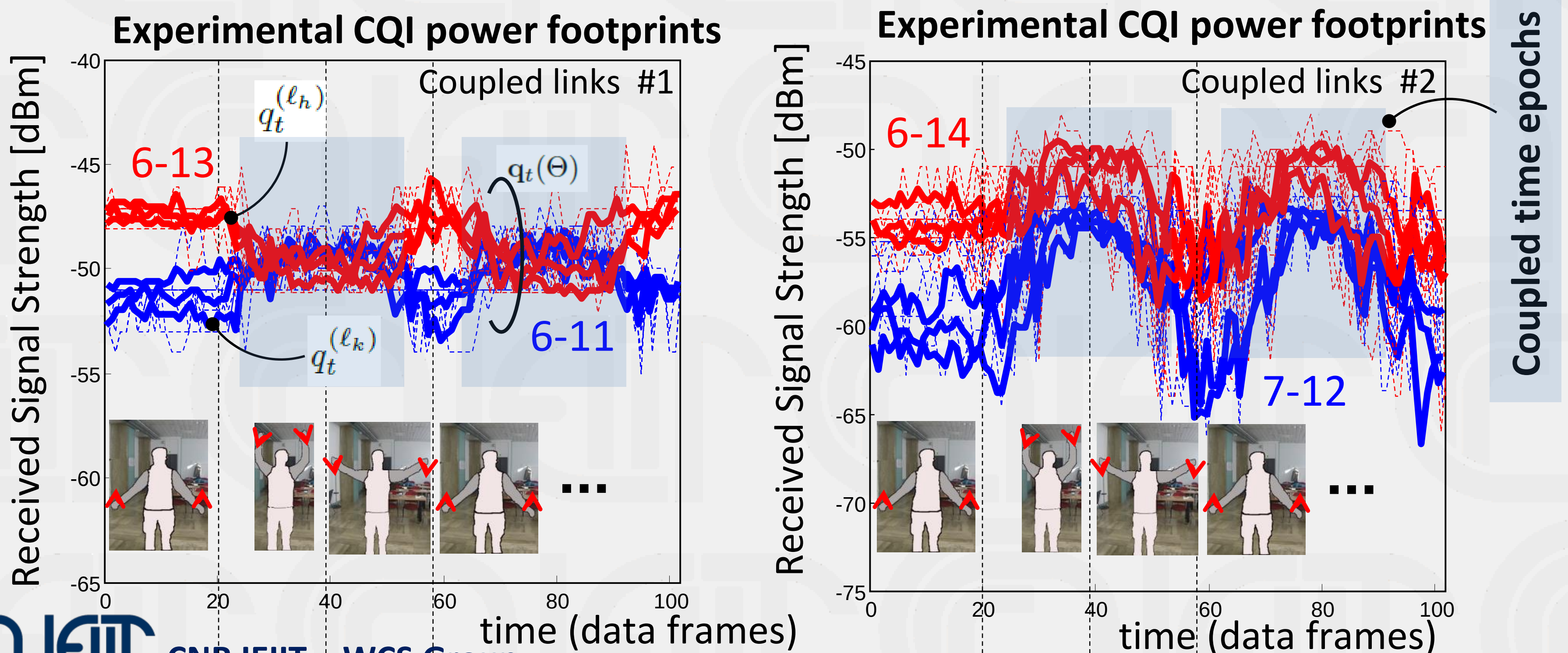
Hidden **CQI shifts** take the role of "link states" and are represented in terms of a set of <u>mutually coupled random variables</u>

$$\mathbf{q}_t(\Theta) \;=\; \left[q_t^{(\ell_1)}, \ldots, q_t^{(\ell_N)}\right]$$

- The collection of temporal shifts in $1 \leq t \leq \mathrm{T}$ is the **CQI profile**.
- Interactions of CQI shits among co-located links <u>are time-varying </u>(therefore standard HMM not effective)
- Focus of this paper is on interactions <u>among link pairs</u>

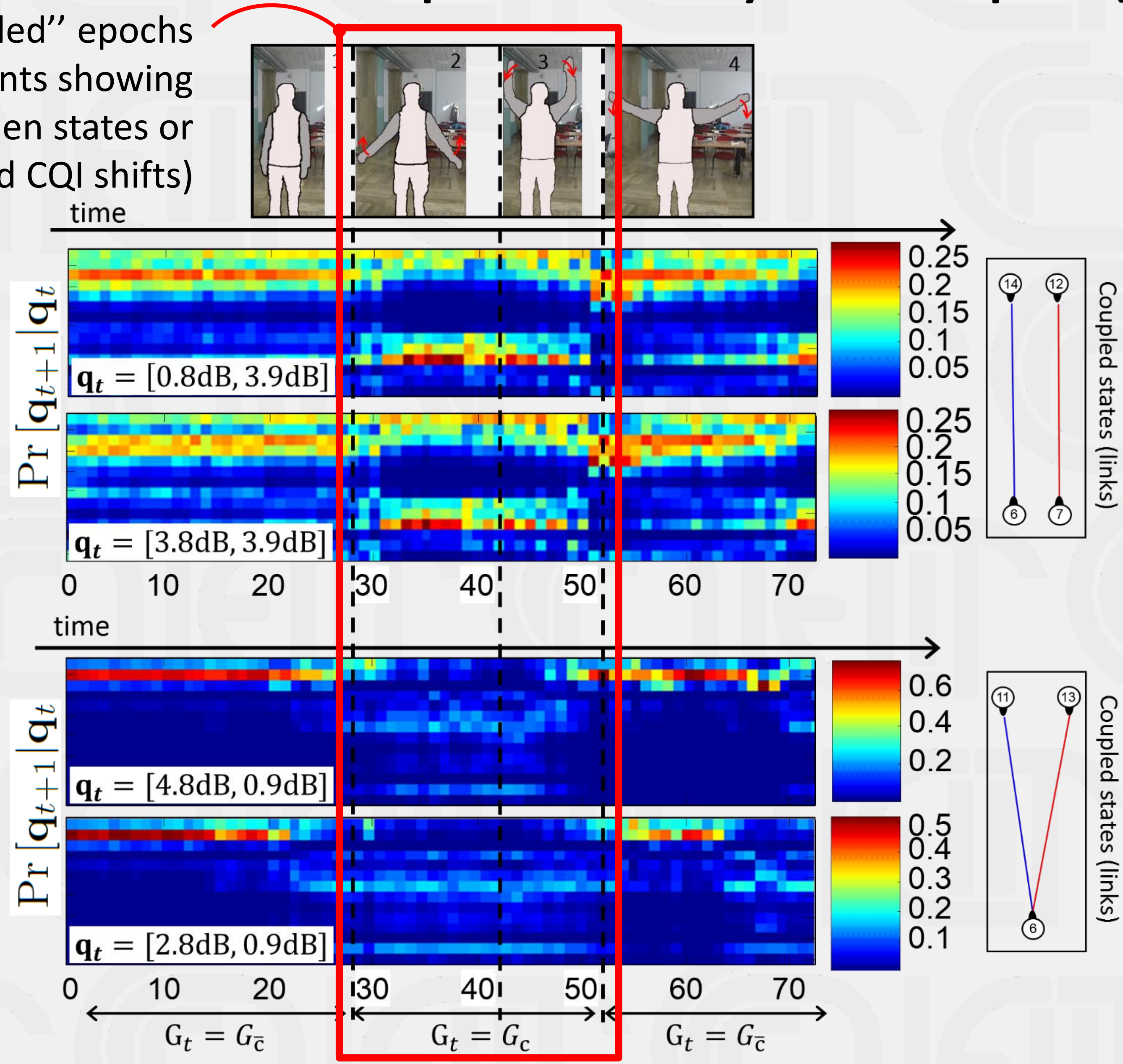
# Gesture recognition: example with RSS data



# State transition probability: example (2)

"coupled" epochs (time instants showing coupled hidden states or correlated CQI shifts)



# Time-varying DBN model for a selected link pair

## 1. Prior network structure $\mathrm{G}_0$

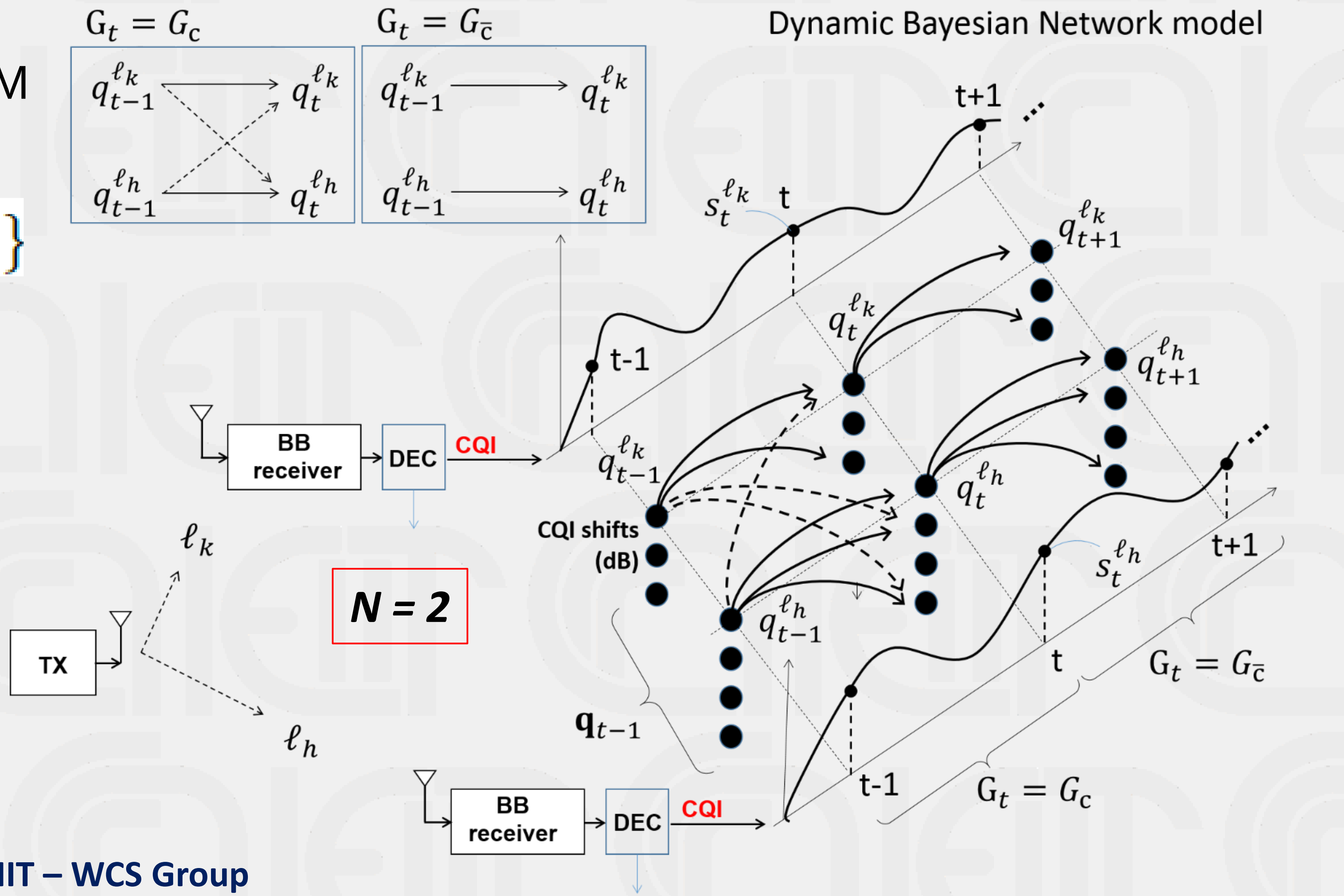
$$\Pr\left[\mathbf{q}_0|\mathrm{G}_0\right] = \prod_{k=1}^{N} \Pr\left[q_0^{(\ell_k)}|\mathrm{G}_0\left(q_0^{(\ell_k)}\right)\right]$$

## 2. Transition network graph $\mathrm{G}_t$

$$\Pr\left[\mathbf{q}_{t+1}|\mathbf{q}_t,\mathrm{G}_t\right] = \prod_{k=1}^{N} \Pr\left[q_{t+1}^{(\ell_k)}|\mathrm{G}_t\left(q_t^{(\ell_k)}\right)\right]$$

## 3. Coupled HM model:

$$\mathrm{G}_t := \{G_{\mathrm{c}}, G_{\bar{\mathrm{c}}}\}$$



# Time-varying DBN model for a selected link pair: (2)

Selected pairs of co-located links $\mathcal{C}_{k,h} = (\ell_k, \ell_h)$ can interact (for some relevant time epochs $t$) by mutually influencing each other future states

Time-varying transition network for paired links can switch among two "topologies" $\forall t$ as

$$\mathrm{G}_t := \{G_{\mathrm{c}}, G_{\bar{\mathrm{c}}}\}$$

to represent coupled and uncoupled configurations at epoch $t$.

The (binary) sequence

$$\mathrm{G}_{1:\mathrm{T}-1} := \{\mathrm{G}_1, \ldots, \mathrm{G}_{\mathrm{T}-1}\}$$

thus rules the time-varying coupling of the embedded CQI shifts.

# Classification and detection of gestures

Likelihood evaluation for activity classification is based on the forward-backward algorithm (frontier method): the joint probability for a selected link pair:

$$\Pr\left[\mathbf{S}_{1:t}, \mathbf{q}_t = \mathbf{q}_{n,q} | \mathbf{G}_{1:t-1}\right] = \alpha_t(\mathbf{q}_{n,q}|\mathbf{G}_{1:t-1})$$

$$\mathbf{q}_{n,q} = \left[q_t^{(\ell_k)} = \mathrm{q}_n, q_t^{(\ell_h)} = \mathrm{q}_q\right]$$

is iteratively evaluated and accounts for the time-varying coupled state sequence $\mathbf{G}_{1:\mathrm{T}-1} := \{\mathbf{G}_1, ..., \mathbf{G}_{\mathrm{T}-1}\}$

$$\alpha_{t+1}\left(\mathbf{q}_{t+1} = \mathbf{q}_{m,p}|\mathbf{G}_{1:t}\right) = \Pr\left(\mathbf{s}_{t+1}|\mathbf{q}_{m,p}\right) \times \times \sum_{\mathbf{q}_{n,q}} \alpha_t\left(\mathbf{q}_{n,q}|\mathbf{G}_{1:t-1}\right) \times \boxed{a_{m|n,q}^{(\ell_k)} a_{p|n,q}^{(\ell_h)}}$$

Time-varying transition network for coupled/uncoupled link pair

Decision on activity $\widehat{\Theta}$: $\Gamma_{\zeta_{k,h}}(\Theta) \geq \tau$

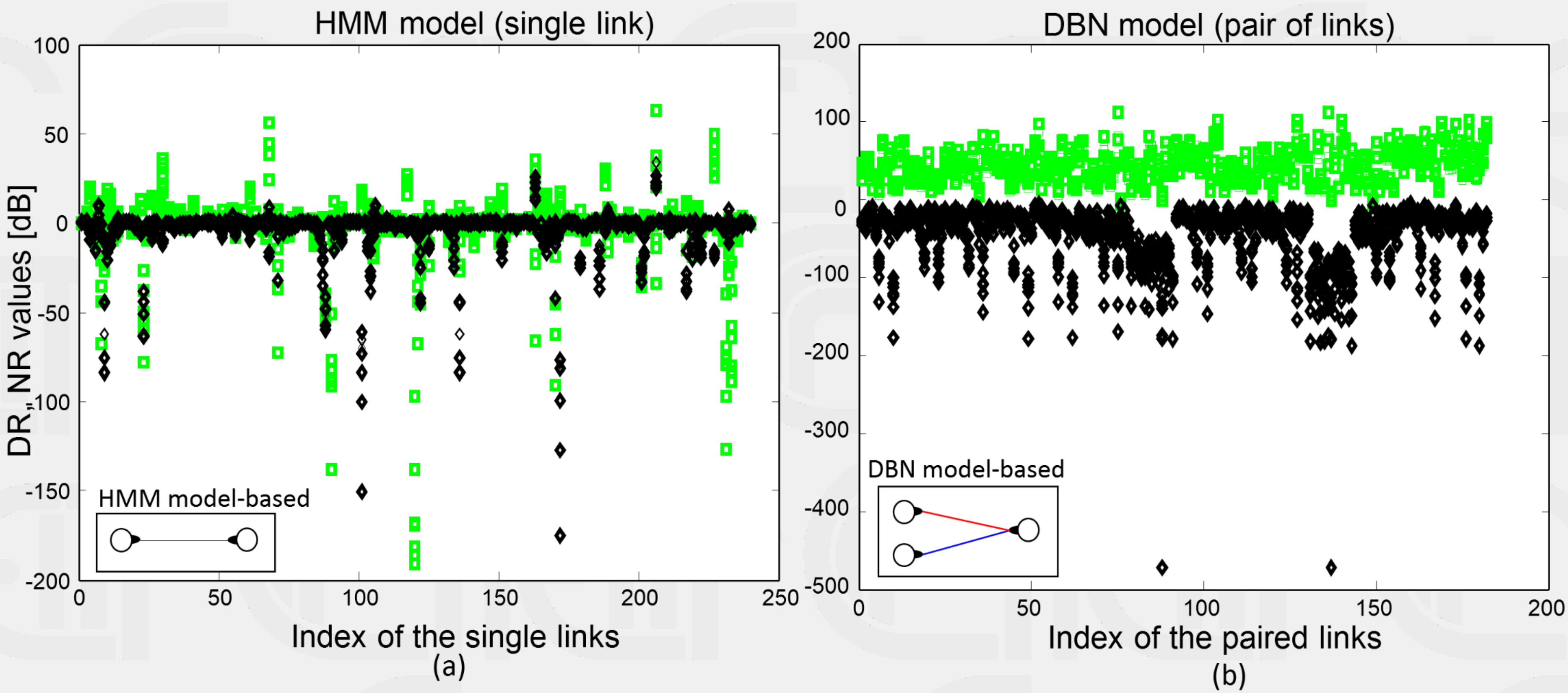
Log-likelihood rate: $\Gamma_{\zeta_{k,h}}(\Theta) = \ln \frac{\Pr[\mathbf{S}_{1:\mathrm{T}}|\lambda(\Theta)]}{\Pr[\mathbf{S}_{1:\mathrm{T}}|\lambda(\emptyset)]}$ $\Pr\left[\mathbf{S}_{1:\mathrm{T}}|\lambda\right] = \sum_{\mathbf{q}_\mathrm{T}} \alpha_\mathrm{T}(\mathbf{q}_\mathrm{T}|\mathbf{G}_{1:\mathrm{T}-1})$

Major voting over selected link pairs

# Gesture/no-gesture detection ratio

DR $\rightarrow$ $$\Gamma_{\zeta_{k,h}}(\Theta) = \ln \frac{\Pr[\boxed{\mathbf{S}_{1:T}}|\lambda(\Theta)]}{\Pr[\boxed{\mathbf{S}_{1:T}}|\lambda(\emptyset)]} \qquad (\mathbf{S}_{1:T} = \mathbf{S}_{1:T}(\Theta))$$

NR $\rightarrow$ $$\Gamma_{\zeta_{k,h}}(\Theta) = \ln \frac{\Pr[\boxed{\mathbf{S}_{1:T}}|\lambda(\Theta)]}{\Pr[\boxed{\mathbf{S}_{1:T}}|\lambda(\emptyset)]} \qquad (\mathbf{S}_{1:T} = \mathbf{S}_{1:T}(\emptyset))$$



# TV-DBN model learning (overview)

- Link pairs chosen during calibration
- Expectation-Maximization (EM) algorithm for DBN model parameters

Starting from a DBN model estimate $\lambda^{(j)}$ at iteration $j$, it re-estimates the DBN parameter set $\lambda^{(j+1)}$ given a new observed sequence $\bar{\mathbf{S}}_{1:\mathrm{T}}^{(j+1)}$ such that

$$\Pr\left[\bar{\mathbf{S}}_{1:\mathrm{T}}^{(j+1)}|\lambda^{(j+1)}\right] \geq \Pr\left[\bar{\mathbf{S}}_{1:\mathrm{T}}^{(j+1)}|\lambda^{(j)}\right]$$

- Re-estimation of intial states and observation probability (standard Baum-Welch)
- Re-estimation of time-varying transition network: $a_{m|n,q}^{(\ell_k)}(j+1),\ a_{p|n,q}^{(\ell_h)}(j+1)$

e.g. for coupled epochs:

$$a_{m|n,q}^{(\ell_k)}(j+1) = \frac{\sum_{\forall t|\mathrm{G}_t=G_c}\sum_p \xi_{t,c}\left(\mathbf{q}_{m,p},\mathbf{q}_{n,q}\right)}{\sum_{\forall t|\mathrm{G}_t=G_c}\pi_{t,c}\left(\mathbf{q}_{n,q}\right)}$$

$$\xi_{t,c} = \Pr\left[\mathbf{q}_{t+1},\mathbf{q}_t|\bar{\mathbf{s}}_t,\mathrm{G}_t^{(j+1)}=G_c,\lambda^{(j)}\right]$$

- Re-estimation of the binary sequence of coupled/uncoupled epochs:

$$\mathbf{G}_{1:\mathrm{T}-1}^{(j+1)} = \arg\min_{\forall \mathbf{G}_{1:\mathrm{T}-1}}\left\|\mathbf{G}_{1:\mathrm{T}-1}-\mathbf{G}_{1:\mathrm{T}-1}^{(j)}\right\|$$
$$\text{s.t.}\Pr\left[\bar{\mathbf{S}}_{1:\mathrm{T}}^{(j+1)}|\lambda^{(j+1)}\right] \geq \Pr\left[\bar{\mathbf{S}}_{1:\mathrm{T}}^{(j+1)}|\lambda^{(j)}\right]$$

Re-estimation of TV transition network

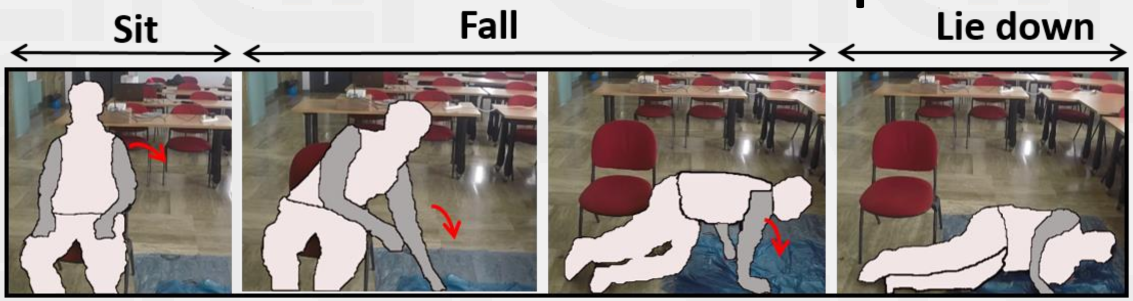
$a_{m|n,q}^{(\ell_k)}(j+1)$

$a_{p|n,q}^{(\ell_h)}(j+1)$

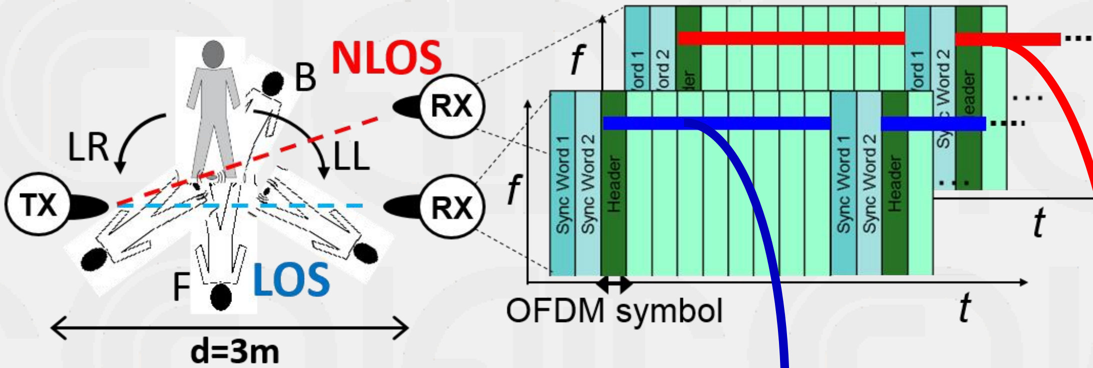
$\mathbf{G}_{1:\mathrm{T}-1}^{(j+1)}$

Re-estimation of binary sequence

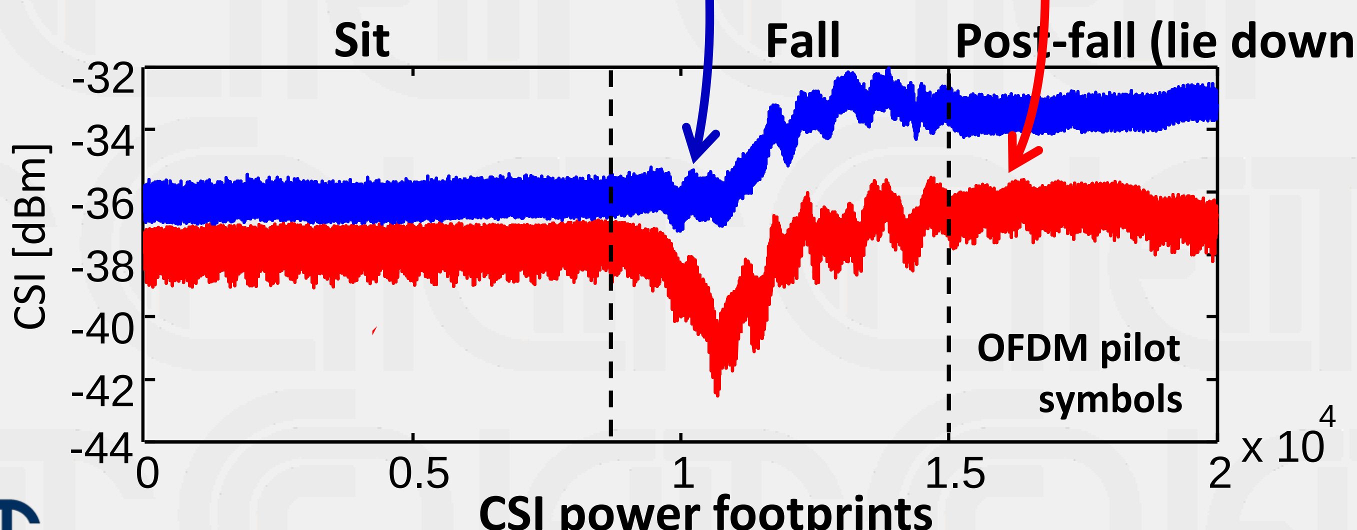
# Fall detection: example with CSI data




**Fall: forward (F), backwards (B), lateral left (LL) and right (LR).**

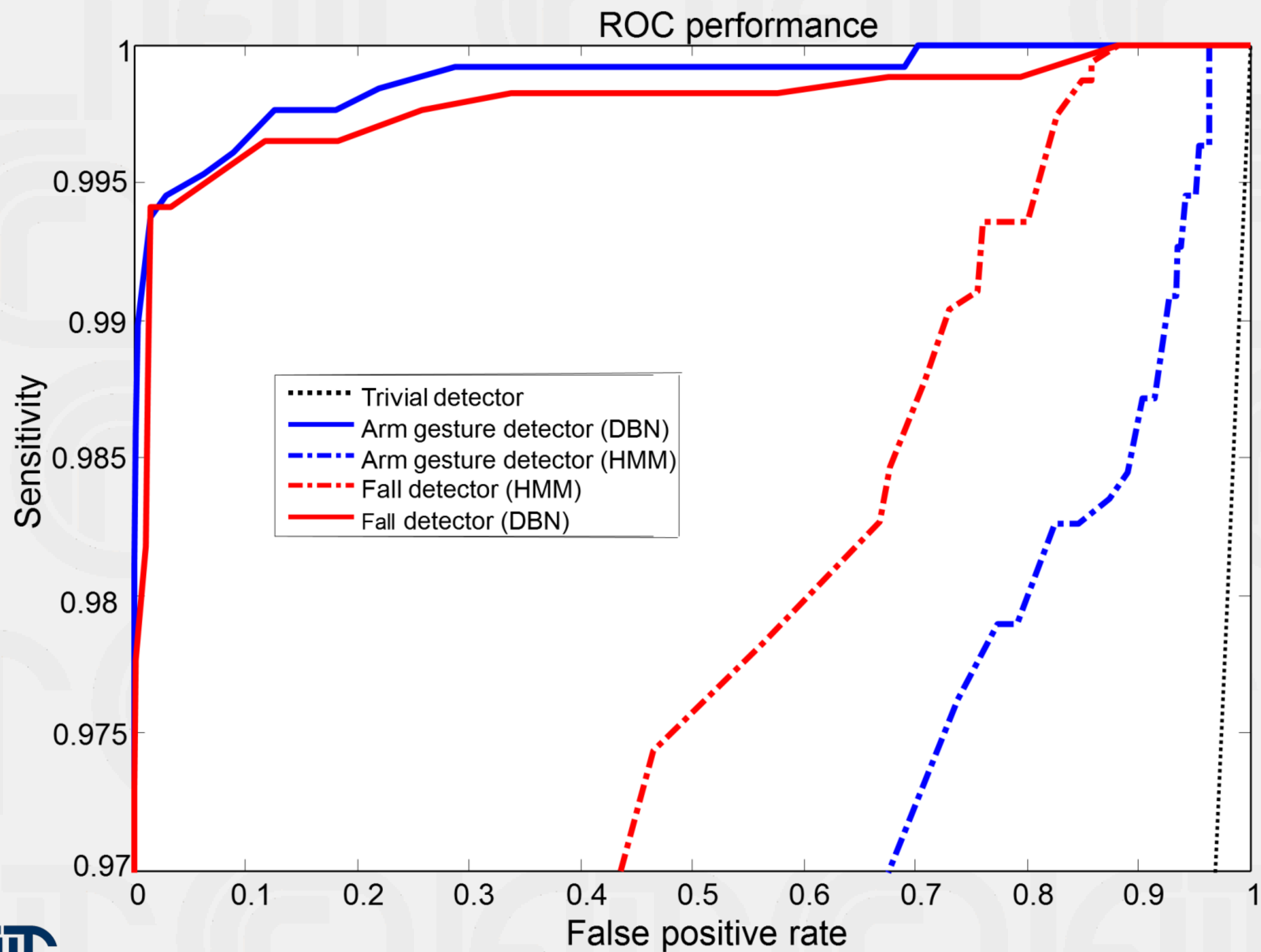



- SDR (USRP N210) device emulating OFDM transceiver
- 1 TX - 2 RXs
- OFDM implementation: 2.6 GHz, 64 sub-carriers, 4 pilots, with 16 payload symbols per frame.
- CQI profiles extracted from one pilot sub-carrier (for both the receivers)



# Validation: ROC curves

## Validation: ROC for multi-link arm gesture and fall detection



# Conclusions

DBN-based techniques proposed for device-free radio vision systems

A time-varying DBN model describes the human-induced CQI footprints, and account for non-stationarity and spatial (link-wise) correlation (coupling) of the channel response over multiple links.

Model is validated or human body motion recognition through extensive experimental RF measurements, focusing on arm gesture recognition and fall detection

Validation of detection performance is analyzed in terms of sensitivity and false positive rates (ROC)