MOBILE DISTRIBUTED COMPRESSIVE SENSING FOR SPECTRUM SENSING

Veria Havary-Nassab*, Shahrokh Valaee*, Shahram Shahbazpanahi[†]

 * Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, ON, Canada, M5S 3G4
 [†]Department of Electrical, Computer, and Software Engineering, University of Ontario Institute of Technology, 2000 Simcoe Street North, Oshawa, ON, Canada, L1H 7K4

ABSTRACT

This paper studies the effect of mobility on the sensing performance of a cognitive radio network with mobile nodes. The secondary nodes sense the spectrum using a distributed compressive sensing approach to detect the available channels. Distributed compressive sensing is suggested to reduce the number of samples by exploiting correlation between the samples. Channel occupancy at the two nodes will be jointly estimated and a channel available at the location of both nodes is chosen for communication. We show that mobility can be exploited to further decrease the number of samples by increasing the average level of correlation among the sensed samples over time.

1. INTRODUCTION

In [1], we proposed the novel method of wide-band compressive estimation for spectrum sensing to address the problem of channel occupancy detection using the sparsity of the channel utilization. As a cognitive radio is interested only in detecting the channel occupancy, this approach estimates the energy of each channel without the need for reconstructing the whole spectrum of the received signal. This method achieves advantages both in complexity and sampling requirements compared to similar approaches such as [2,3].

The approach proposed in [1] uses compressive sensing and requires that all nodes separately sample the spectrum and detect the available channels at their respective locations. Such channel occupancy patterns are then used to find a channel available at the location of all nodes. While this approach benefits from the sparsity of the channel occupancy to reduce the sampling requirements, it does not address the correlation between the samples acquired by the communicating nodes. Such a correlation can indeed help to further reduce the required number of samples per node, leading to a gain in number of samples. The sensing gain is defined as the amount by which the number of required samples for successful estimation is reduced. In this paper, we first propose a joint channel occupancy reconstruction approach to decrease the number of samples each node needs to acquire for successful estimation. Then we show that mo*bility increases the sensing gain* by increasing the probability of high correlation between the nodes' samples.

In distributed compressive sensing, spatial correlation between samples taken by nearby sensors is used to reduce the number of samples required to reconstruct the signal [4, 5], [6]. In a mobile wireless network, the sensing and detecting procedures need to be repeated periodically to address the fact that mobility will change the topology of the network. While this might suggest that mobility degrades the performance of the estimation, properly exploiting mobility has been shown to improve the performance of wireless sensor networks [7, 8]. It has been shown that for a network with growing number of nodes in a limited area and in the presence of interference, mobility can enhance the aggregate capacity of a mobile ad-hoc network.

In [9], we studied the effect of mobility on the average number of opportunities a mobile sensor node will have in a limited time, to communicate with one of several mobile sink nodes in its vicinity. We proved that higher mobility of sensor nodes leads to higher opportunities for sensors to get in the coverage of a sink node. Using a similar approach toward node mobility, we herein show that mobility also benefits the proposed distributed compressive sensing based algorithm by allowing higher average correlation between the sensed signals, thereby leading to a smaller number of needed samples.

2. PROBLEM SETUP

Consider a cognitive radio scenario, where n primary users are communicating over some of the N available channels in a network. Some secondary nodes are also trying to use the spectrum assigned to the primary users opportunistically to exchange information. Assume that two secondary nodes decide to communicate over one of the available channels. The two secondary nodes sense the environment periodically, at time instants $\{t_i = i\Delta t\}_{i=1}^M$ to find available channels for communication over the $T = M\Delta$ seconds of the network activity. Assume that at each of these instants, a primary node is using its dedicated channel with probability $p \ll 1$. Also assume that the average number of occupied channels at each time instant, $S \triangleq np \ll N$ meaning that the channel occupancy is sparse.

The two secondary nodes, which intend to share information, sense the spectrum in a collaborative manner to exploit the correlation between the channel occupancy patterns at the location of the nodes. Denoting the vector of samples taken by a secondary node by \mathbf{y} and the vector of energies in each of the N channels of the spectrum at the location of that node by \mathbf{e} , we can write

$$\mathbf{y} = \mathbf{\Phi} \mathbf{e} \tag{1}$$

where Φ is the sampling matrix defined by the transfer functions of the filters embedded in the secondary nodes [1]. Assuming that the channel is mostly unused (i.e. e is sparse), a compressive sensing reconstruction algorithm (such as the ℓ_1 norm minimization) can reconstruct e from y.

Now assume that two secondary nodes intend to communicate and therefore need to estimate the channel occupancies at their locations. Using a narrow-band control channel, the two SUs decide on one node to do the joint estimation and the other node shares its sensing with the estimator node. When the two nodes sense the spectrum, the occupied channels they detect can come from two categories of primary nodes. First, the primary nodes in the interference range of *both* nodes will introduce correlated energy levels in both nodes in channels they occupy (common component). Second category are primary nodes that lie in the interference range of one of the two secondary nodes (innovation). Denoting the vector of channel energies at the location of the two secondary nodes by \mathbf{e}_1 and \mathbf{e}_2 , using the joint sparsity model 1 (JSM-1) [4] we can write

$$y_{1} = \Phi e_{1} \quad e_{1} = w + z_{1}$$

$$y_{2} = \Phi e_{2} \quad e_{2} = w + z_{2}$$

$$|w||_{0} = L , ||z_{1}||_{0} = L_{1} , ||z_{2}||_{0} = L_{2}$$
(2)

Here, the length-N and L-sparse energy vector w represents the common component of e_1 and e_2 related to the L nodes that occupy the channels at the location of both nodes. The vectors z_1 and z_2 stand for the L_1 and L_2 nodes that occupy the channels only at node 1 and only at node 2, respectively.

Using this model, the estimator node solves a linear programming optimization problem to simultaneously reconstruct e_1 and e_2 using $C(L + L_1 + L_2)$ samples where C is the coherence coefficient representing the correlation between the sensing matrix and the sparsity basis. This is done through choosing

$$\mathbf{e}_{T} = \begin{bmatrix} \mathbf{w} \\ \mathbf{z}_{1} \\ \mathbf{z}_{2} \end{bmatrix} \mathbf{y}_{T} = \begin{bmatrix} \mathbf{y}_{1} \\ \mathbf{y}_{2} \end{bmatrix} \mathbf{\Phi}_{T} = \begin{bmatrix} \mathbf{\Phi} & \mathbf{\Phi} & 0 \\ \mathbf{\Phi} & 0 & \mathbf{\Phi} \end{bmatrix}.$$
(3)

and solving

$$\mathbf{e}_{T} = \arg\min_{\mathbf{e}_{T}} \|\mathbf{e}_{T}\|_{1}$$

Subject to $\mathbf{y}_{T} = \mathbf{\Phi}_{T}\mathbf{e}_{T}.$ (4)

The channel occupancy vectors will then be used to choose an empty channel for communication between the secondary nodes. The vector \mathbf{e}_T has a maximum of $L + L_1 + L_2$ non-zero components and therefore, a total of $C(L + L_1 + L_2)$ samples are sufficient to reconstruct \mathbf{e}_T . If node 1 and node 2 decide to reconstruct the channel energy vectors \mathbf{e}_1 and \mathbf{e}_2 separately, they would need $C(L+L_1)$ and $C(L + L_2)$ samples, respectively. This shows that if the two nodes share their sample vectors $(\mathbf{y}_1 \text{ and } \mathbf{y}_2)$ and reconstruct the channel energies jointly, each node can take *fewer* samples compared to the separate reconstruction case. The total sensing gain (amount of decrease in required number of samples) is therefore CL.

3. EFFECT OF MOBILITY ON THE PERFORMANCE OF THE PROPOSED METHOD

Define $L(t_i)$ to be the number of channels detected to be occupied in *both* nodes at time t_i due to the presence of primary nodes interfering with both secondary nodes.

The two secondary nodes use joint channel occupancy reconstruction with linear programming as discussed in Section 2. They fix the number of samples each node takes based on a nominal prior for the value of L named $L_{\rm P}$ to achieve a desirable probability of detection P_D . In other words, the nodes not only assume that the channel occupancy is sparse, but they also consider a level of correlation between their samples and adjust their sampling requirements based on this assumption. Using the results of Section 2, each node acquires $K_1 = K_2 = \frac{1}{2}C(L_{\rm P} + L_1 + L_2)$ samples based on the nominal $L_{\rm P}$ and an average of $L_1 = L_2 = S - L_{\rm P}$ for the sparsity level of the innovation signals (number of primary nodes interfering with one of the secondary nodes only). However, the instantaneous values of L, L_1 and L_2 are random due to the random mobility of the nodes.

The success of the joint channel occupancy reconstruction at each time instant t_i will therefore depend on the actual $L(t_i)$. If the level of correlation between the sensed signals is higher than the prior assumption, $L(t_i) \ge L_{\rm P}$, the reconstruction will be successful with a probability at least equal to the desired P_D because the number of samples taken is higher than the number of required samples. If $L(t_i) < L_{\rm P}$, the detection has a failure probability higher than $1 - P_D$ at $t = t_i$ and the communication might cause interference to the primary network at the interval $[t_i \ t_{i+1}]$. Based on the scenario and configurations of the problem (number of nodes, interference threshold, the mobility of the network), a proper $L_{\rm P}$ can be obtained that minimizes the required number of samples while keeping the probability of detection at the desired level.

The effect of mobility is therefore studied by investigating how mobility of the nodes changes the average value of the parameter $L(t_i)$. If mobility is shown to increase $L(t_i)$ on average and over the course of the communication (over the M time instants $\{t_i = i\Delta t\}_{i=1}^M$), it can either increase the probability of successful detection for a fixed L_P , or reduce the L_P needed to achieve a certain P_D . We assume that the secondary nodes are mobile in a square grid of total area A_T . The n primary nodes are assumed to have a uniform distribution of the same grid. We further adopt a two dimensional Wiener mobility model for the secondary nodes similar to the one adopted and analyzed in [9]. Using Wiener process (or Levy process in general) and random walk is a common practise for modelling mobility in wireless networks for several reasons including tractability, level of uncertainty in direction of the nodes' motion and independent increment property [10–13].

In this model, each of the x and y coordinates of a secondary node will be formulated by an independent Wiener random process given by

$$X_i(t) = X_i(0) + \sigma V_{i,1}(t)$$
 (5)

$$Y_i(t) = Y_i(0) + \sigma V_{i,2}(t), \quad i = 1, 2.$$
 (6)

Here $\{X_i(0)\}_{i=1,2}$ and $\{\mathbf{y}_i(0)\}_{i=1,2}$ are the random variables of the coordinates of the two secondary nodes at t = 0, whereas σ is the common *mobility parameter* of the secondary nodes. The parameter σ controls the level of mobility each node possesses. Smaller values of σ correspond to lower levels of mobility and larger values of σ represent nodes that are more mobile. The random processes $V_{i,1}(t)$ and $V_{i,2}(t)$ are statistically independent standard Wiener processes defined as $V_i(t) \sim \mathcal{N}(0, t), V_i(0) = 0, i = 1, 2$. Here, \sim is used to show the distribution of a random variable and $\mathcal{N}(a, b)$ refers to a normal distribution with mean a and variance b.

Analysis

In the following analysis, we first show that the average number of primary nodes creating correlated samples in the two secondary nodes decreases with the distance of the two secondary nodes. Assuming that the primary nodes are using a constant transmit power, the interference threshold can be translated into a frequency reuse distance. A secondary node will interfere with the primary node if their distance is smaller than D.

Lemma 3.1 For $d \in [0, 2D]$,

$$E\{L(t_i)\} = p\lambda\left(2D^2\arccos\frac{d}{2D} - dD\sqrt{1 - (\frac{d}{2D})^2}\right)$$
(7)



Fig. 1. A simple method to bound $Pr\{d(t_i) < d\}$. Here a = d, and $b = d/\sqrt{2}$

where λ is the density of the number of primary nodes per unit surface.

Proof At any time t_i the average number of active nodes within a region with area A can be expressed as $N(A) = p\lambda A$. The average value of the parameter $L(t_i)$ for two secondary nodes with distance d is therefore equal to $N(A_C(d))$ where $A_C(d)$ is the region shared between two circles of radius D and center distance d. It is easy to see that

$$A_C(d) = 2D^2 \arccos \frac{d}{2D} - dD\sqrt{1 - (\frac{d}{2D})^2}$$
 for $d \in [0, 2D]$.

Corollary 3.2 *Expected value of* $L(t_i)$ *is a decreasing function of the distance between the two nodes, d.*

Proof The derivative of $L(t_i)$ with respect to the nodes' distance d is

$$-p\lambda\left(\frac{D-d/2}{\sqrt{1-(\frac{d}{2D})^2}} + D\sqrt{1-(\frac{d}{2D})^2}\right)$$
(8)

which is clearly negative for 0 < d < 2D.

Denote by $I(t_i)$ the indicator random variable for the event $L(t_i) \ge L_{\rm P}$, or equivalently, $I(t_i) = 1$ if $L(t_i) \ge L_{\rm P}$ and 0 otherwise. The random sequence $I(t_i)$ is therefore an indicator of a successful detection with probability P_D . If $L(t_i) \ge L_{\rm P}$, the number of samples that the two secondary nodes acquire will be sufficient for a successful reconstruction of the energy vectors with probability P_D . Let us define

$$N_s = \sum_{i=1}^{M} I(t_i) \tag{9}$$

to be the random variable of the total number of time instants in which the $L(t_i) \ge L_{\rm P}$, or equivalently, the number of time instants at which the nodes have successfully reconstructed the channel occupancy with the desired P_D . In the following theorem, we prove that the higher the mobility of the nodes, the larger the average of N_s . In other words, mobility can contribute to the performance of our distributed compressive estimation algorithm by providing a larger correlation among the signals sensed by the mobile sensors. Such increase can be used in two fashions. If the same value for $L_{\rm P}$ is adopted, the increase results in a higher reliability communication within the secondary nodes. However, if the current success rate in communication is acceptable, a more mobile network can adopt a larger $L_{\rm P}$ and therefore take fewer samples per node.

Theorem 3.3 The probability $Pr\{L(t_i) > L_P\}$ and therefore the average number of successful detections in M time instants, $E\{N_s\}$, are increasing functions of σ , the mobility parameter of the network.

Proof We first define $d_{\rm P}$ to be the distance between the two secondary nodes corresponding to $L = L_{\rm P}$ primary nodes interfering with both sensor nodes. In other words, assume that $N(A_C(d_{\rm P})) = L_{\rm P}$. This means that, when the distance between the two nodes is $d_{\rm P}$, the average number of channels occupied at the two nodes due to primary nodes interfering both secondary nodes is $L_{\rm P}$. It is now easy to see that

$$\Pr\{L(t_i) > L_{\mathrm{P}}\} \equiv \Pr\{N(A_C(d(t_i))) > N(A_C(d_{\mathrm{P}}))\}$$
$$\equiv \Pr\{A_C(d(t_i)) > A_C(d_{\mathrm{P}})\}$$
(10)
$$\equiv \Pr\{d(t_i) < d_{\mathrm{P}}\}$$
(11)

where we have replaced d with $d(t_i)$ to explicitly denote the time dependence of the distance between the two nodes. The equivalence in (10) is a result of N(A) being an increasing function of A and the second equivalence, (11) is a result of $A_C(d)$ being a decreasing function of d. Assuming that the two nodes are mobile based on the Wiener process introduced in (5) and (6), we can write

$$\Pr\{d(t_i) < d_{\rm P}\} = \Pr\{[X_1(t_i) - X_2(t_i)]^2 + [Y_1(t_i) - Y_2(t_i)]^2 < d_{\rm P}^2\}.$$
 (12)

Figure 1 suggests a very simple approach to bounding $\Pr\{d(t_i) < d_P\}$. Based on this diagram, we can see that $\Pr\{\mathcal{A}\} < \Pr\{d(t_i) < d_P\} < \Pr\{\mathcal{B}\}$ where the events \mathcal{A} and \mathcal{B} are defined as

$$\mathcal{A} = \{\Delta X(t_i) < d_{\mathrm{P}}/\sqrt{2}\} \text{ and } \{\Delta Y(t_i) < d_{\mathrm{P}}/\sqrt{2}\}$$
$$\mathcal{B} = \{\Delta X(t_i) < d_{\mathrm{P}}\} \text{ and } \{\Delta Y(t_i) < d_{\mathrm{P}}\}.$$

Here $\Delta X(t_i) = |X_1(t_i) - X_2(t_i)|$ and $\Delta Y(t_i) = |Y_1(t_i) - Y_2(t_i)|$. Using the assumption that the random processes representing the x and y coordinates of the nodes are independent, we can conclude that the lower bound and the upper bound of $\Pr\{d(t_i) < d_P\}$ can be respectively written as

$$\Pr{\mathcal{A}} = \Pr{dX(t_i) < \frac{d_{\mathrm{P}}}{\sqrt{2}}} \Pr{dY(t_i) < \frac{d_{\mathrm{P}}}{\sqrt{2}}}$$
$$\Pr{\mathcal{B}} = \Pr{dX(t_i) < d_{\mathrm{P}}} \Pr{dY(t_i) < d_{\mathrm{P}}}.$$

In other words, the events whose probabilities establish a lower bound and an upper bound for $\Pr\{d(t_i) < d_{\rm P}\}$ are cross sections of two independent events and therefore the bounding probabilities can be written as the multiplication of the probabilities of the underlying independent events. Each of the two events appearing in the lower bound and the upper bound of $Pr\{d(t_i) < d_P\}$ corresponds to the event that the distance between two one-dimensional Wiener processes does not exceed a fixed value ($d_{\rm P}$ or $d_{\rm P}/\sqrt{2}$). In [14], we have proven that the probability of such events are increasing functions of the mobility parameters of the network nodes, if within the observation time T, the secondary nodes do not *leave* the area covered by the primary nodes. This implies that both the lower bound and the upper bound of $Pr\{d(t_i) < d_P\}$ are increasing functions of the mobility parameter. Therefore $Pr\{d(t_i) < d_P\}$ and its equivalent $\Pr\{L(t_i) > L_{\rm P}\}$ are increasing functions of the mobility parameter.



Fig. 2. Separate reconstruction vs. joint reconstruction



Fig. 3. Effect of mobility on the probability of successful reconstruction

4. SIMULATIONS

Figure 2 illustrates the result of a numeric simulation to determine the number of required samples to successfully detect the vector of channel energies in three different scenarios:

- 1. Separate reconstruction
- 2. Joint reconstruction, L = 2, $L_1 = L_2 = 4$
- 3. Joint reconstruction, L = 4, $L_1 = L_2 = 2$

Each node senses the spectrum and reconstruct the energy vector separately (in scenario 1) or jointly (in scenarios 2,3) with different number of samples (5 – 15 samples). As seen in Fig. 2, with 12 samples and using separate reconstruction, each node can detect the channel occupancy with the probability of 0.9. In scenario 2, joint reconstruction method discussed in Section 2 is used to retrieve the channel energy vectors \mathbf{e}_1 and \mathbf{e}_2 . As seen in Fig. 2, if each of the nodes takes 10 samples, this approach can reconstruct the channel occupancy with probability more than 90%. Fig. 2 suggests the required samples per node in scenario 3 is just 8. To address the effect of mobility, we have considered a scenario where in a 50x50 grid, 20 primary nodes and two secondary nodes are moving randomly based on a Gaussian random walk model with step size σ . A channel is



Fig. 4. Required number of samples per node to achieve $P_D=90\%$ vs. mobility parameter, σ obtained by simulation and predicted by analysis (details of analysis not included in this paper)

occupied within D = 10 meters of the primary users communicating on that channel. A total of 20 channels are available for nodes to communicate and the probability of any primary node using its assigned channel is p = 6/20. The two nodes intending to communicate sense the spectrum every $\Delta t = 5$ seconds. At each time instant, the two secondary nodes use 5 to 15 samples per node to jointly reconstruct the channel occupancy pattern based on the approach discussed in Sections 3 and 2. The network is simulated over 100 seconds and repeated several times to average the random motion of the nodes. The whole process is repeated with nodes having three different step sizes, $\sigma = 2, 4, 6$. As seen in Figure 3, the more mobile the nodes are, the higher the correlation between the sensed signals will be and therefore, even fewer number of samples can be used to successfully reconstruct the channel energy vectors. The gain observed in Fig. 3 is due to the fact that in a network with a larger mobility, the communicating nodes will have a higher probability of being in a closer distance of each other over the fixed time instants $t_i, i = 2, 3, \dots, M$ leading to a larger average of $L(t_i)$ and resulting in a higher probability for the success of the joint reconstruction algorithm using smaller number of samples. In Fig. 4 the results of the mobile node simulation has been used to plot the number of samples required to achieve $P_D=90\%$ versus the mobility parameter of the network nodes. The $\sigma = 0$ case corresponds to separate detection (not addressing correlation) that also appears in Fig. 3. As seen in Fig. 4, when the nodes in a network become more mobile, they can take fewer samples and still reconstruct the phenomenon of interest with the same success probability using the proposed joint reconstruction scheme.

5. CONCLUSION

A distributed compressive sensing based approach has been proposed for spectrum sensing in a mobile ad-hoc cognitive radio network. Exploiting the correlation between samples of close-by nodes, the proposed method jointly reconstructs the spectrum occupancy at the location of secondary nodes with fewer number of samples. Furthermore, the effect of the mobility of the nodes has been studied using a Wiener process model. It has been proven that by mobility can be exploited to further reduce the number of samples, thereby introducing a sensing gain.

6. REFERENCES

- V. Havary-Nassab, S. Hassan, and S. Valaee, "Compressive detection for wide-band spectrum sensing," in *IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), 2010, 2010, pp. 3094–3097.
- [2] Y. L. Polo, Y. Wang, A. Pandharipande, and G. Leus, "Compressive wide-band spectrum sensing," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, *ICASSP '09*, Taipei, Apr. 2009, pp. 2337–2340.
- [3] Z. Tian and G. B. Giannakis, "Compressed sensing for wideband cognitive radios," in *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '07*, vol. 4, Honolulu, HI, Apr. 2007, pp. 1357–1360.
- [4] M. Duarte, S. Sarvotham, D. Baron, M. Wakin, and R. Baraniuk, "Distributed compressed sensing of jointly sparse signals," in *Thirty-Ninth Asilomar Conference on Signals, Systems and Computers*, 2005, pp. 1537–1541.
- [5] D. Baron, M. F. Duarte, S. Sarvotham, M. B. Wakin, and R. G. Baraniuk, "An information-theoretic approach to distributed compressed sensing," in 43rd Allerton Conference on Communication, Control, and Computing, Sep. 2005.
- [6] A. Makhzani and S. Valaee, "Reconstruction of a generalized joint sparsity model using principal component analysis," in *Computational Advances in Multi-Sensor Adaptive Processing* (CAMSAP), 2011 4th IEEE International Workshop on, 2011, pp. 269–272.
- [7] S. Diggavi, M. Grossglauser, and D. Tse, "Even onedimensional mobility increases the capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 51, 2005.
- [8] J. Mammen and D. Shah, "Throughput and delay in random wireless networks with restricted mobility," *IEEE Transactions* on *Information Theory*, vol. 53, 2007.
- [9] V. Havary-Nassab, S. ShahbazPanahi, and S. Valaee, "Mobility diversity in mobile wireless networks," in *IEEE 22nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, 2011, sept. 2011, pp. 1093 –1097.
- [10] B. Jabbari, Y. Zhou, and F. Hillier, "Random walk modeling of mobility in wireless networks," in *Vehicular Technology Conference*, 1998. VTC 98. 48th IEEE, vol. 1, 1998, pp. 639–643 vol.1.
- [11] A. Giannakos, G. Karagiorgos, and I. Stavrakakis, "A messageoptimal sink mobility model for wireless sensor networks," in *Networks, 2009. ICN '09. Eighth International Conference on*, 2009, pp. 287–291.
- [12] K. Amjad, "Performance analysis of dsr protocol under the influence of rpgm model in mobile ad-hoc networks," in *Distributed Computing Systems Workshops (ICDCSW)*, 2011 31st International Conference on, 2011, pp. 100–104.
- [13] S. Yarkan, A. Maaref, K. H. Teo, and H. Arslan, "Impact of mobility on the behavior of interference in cellular wireless networks," in *Global Telecommunications Conference*, 2008. *IEEE GLOBECOM 2008. IEEE*, 2008, pp. 1–5.
- [14] V. Havary-Nassab, S. Shahbazpanahi, and S. Valaee. (2013) Mobility diversity in mobile wireless networks. [Online]. Available: http://arxiv.org/abs/1310.1833