Estimation of Cellular Wireless User Coordinates via Channel Charting and MUSIC

Amr Aly, *Student Member, IEEE*
CPCC, Department of EECS
University of California, Irvine, CA, USA
alyas@uci.edu

Ender Ayanoglu, *Fellow, IEEE*
CPCC, Department of EECS
University of California, Irvine, CA, USA
ayanoglu@uci.edu

***Abstract*—We present a new way of producing a channel chart for cellular wireless communications in polar coordinates. We estimate the angle of arrival θ and the distance between the base station and the user equipment ρ using the MUSIC algorithm and inverse of the root sum squares of channel coefficients (ISQ) or linear regression (LR). We compare this method with the channel charting algorithms principal component analysis (PCA), Samson's method (SM), and autoencoder (AE). We show that ISQ and LR outperform all three in both performance and complexity. The performance of LR and ISQ are close, with ISQ having less complexity.**

***Index Terms*—Channel charting, user equipment (UE), channel state information (CSI), MUSIC, PCA, SM, AE.**

I. Introduction

In a cellular wireless communication system, a channel chart is created from channel state information (CSI) that preserves the relative distance information between user equipments (UEs) [1], [2]. This chart helps the base station (BS) locate the UEs, which can help in many applications for cellular wireless communications such as handover, cell search, user localization, and more. Previous papers have proposed estimation of a channel chart using Cartesian coordinates. The seminal reference [1] compared three algorithms, namely principal component analysis (PCA), Sammon's mapping (SM), and autoencoder (AE). All of these algorithms try to convert the channel coefficients into the angular domain and then try to extract two features for each UE and these two features will represent the UE location in the channel chart. Reference [3] tries to use the AE algorithm in a supervised fashion by allowing some of the UEs to have global positioning system (GPS) data for exact location and use it to improve the AE learning of the geometry. In this letter we propose to estimate the location of UE in the channel chart in polar coordinates. We first estimate the angle of arrival (AoA) θ using the MUSIC algorithm based on the channel coefficients correlation matrix [4]. For the distance between the BS and the UE ρ, we first sum the power on all antennas for each UE, then we take the inverse of the root of the sums and use it as ρ. As an alternative, we consider linear regression and compare the results of these approaches and those of PCA, SM, and AE. We use the same simulation environment in [1] with slight modifications.

This work is partially supported by NSF grant 2030029.

II. Estimating θ and ρ

In a cellular wireless communication system, estimating θ and ρ can happen concurrently as they do not depend on each other. We will first discuss how to estimate θ by using the MUSIC algorithm. From Fig. 1, we can see that each antenna element in the multiple antenna array of the BS will receive a ray that travels an additional distance $\frac{\lambda}{2}\cos(\theta)$ than the previous element. This means for each antenna element, the incremental phase shift is $e^{j(\pi\cos(\theta))}$. With this shift, we get what is called the steering vector

$$A = (1, e^{j\pi\cos(\theta)}, e^{j\pi 2\cos(\theta)}, \ldots, e^{j\pi(N-1)\cos(\theta)})^T. \quad (1)$$

This steering vector is embedded within the CSI covariance matrix ($\mathbf{R}$) along with noise. If we decompose $\mathbf{R}$ into its eigenvectors and examine the corresponding eigenvalues, we can separate the eigenvectors into a signal subspace S and a noise subspace N. The noise eigenvectors will have very small eigenvalues. The subspaces S and N are orthogonal to each other. Therefore, the dot product of the noise subspace eigenvectors matrix N and the steering vector will be almost zero. We can use this concept to find the correct angle by sweeping θ in the steering vector as illustrated in Algorithm 1, where PMF stands for probability mass function.

We will now discuss how to estimate ρ. The simple channel ray model can be depicted as

$$h = a(d)\, e^{j\left(\frac{2\pi d}{\lambda}+\phi\right)} \text{ where } a(d) \sim d^{-2}, \quad (2)$$

where the first term in channel phase is linearly proportional with the distance d between the transmitter and the receiver and the second term ϕ is a uniformly distributed random

Algorithm 1 MUSIC Procedure

Calculate the CSI covariance matrix $\mathbf{R} = \mathbb{E}[\mathbf{h}\mathbf{h}^H]$
Get the eigenvectors and eigenvalues of $\mathbf{R}$
Separate system subspace S and noise subspace N
for $\theta = 0 : 180$ in increments of 1/2 **do**
  Calculate the steering vector $A(\theta)$
  Calculate the PMF$(\theta) = \frac{1}{\text{Norm}(A(\theta)\cdot N)}$
end for
Search the PMF for a peak and find the corresponding θ

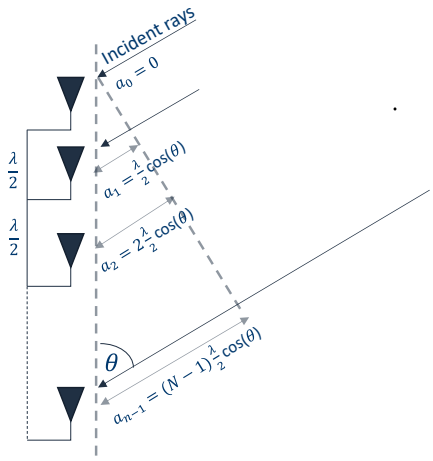

Fig. 1. Angle of arrival (θ) relation with phase.

variable in $[0, 2\pi)$. The channel amplitude is a random variable (Rayleigh(QNLOS) or Rician(QLOS)) which is inversely proportional to the distance square for free space, $\sim d^{-2}$. The number 2 in this expression is called the path loss exponent. For more crowded environments the path loss exponent can be 3 or 4. Our proposal is a rather direct and simple approach. We calculate the square root inverse of the sum of CSI magnitudes for all antennas as

$$\rho = \frac{1}{\sqrt{\sum_{n=0}^{N-1} \text{abs}(h_n)}}. \tag{3}$$

Earlier, we tried a supervised approach by assuming we know the location of 256 (out of 2048) UEs and do a linear regression with the logarithm of the sum of CSI magnitudes for all antennas to find a and b in

$$\rho = aX + b \ , \text{ where } X = \log \sum_{n=0}^{N-1} \text{abs}(h_n). \tag{4}$$

But, as we will show later, the unsupervised performance is almost identical to the linear regression. We will show later that both methods correlate with real ρ linearly.

III. Simulation Environment and Assumptions

In this letter we reused and integrated our algorithm into the simulation environment in [1] so that we can fairly compare the performance improvement. We adopted the simulation parameters in Table III. We assume the antenna elements and the UEs are in the same plane. We present three cases depending on the dimensions of the simulation environment which is 1000m x 500m as in [1]. In the first case, the dimensions are the same as the dimensions in [1]. In the second case we see the performance of reducing this to 50% of the original dimensions. In the last case, the dimensions are reduced to 25% of the original dimensions. We refer to our algorithms as ISQ (inverse square root sum) and LR (linear regression).

TABLE I
Simulation Parameters for the Cellular Wireless Communication System

Parameter	Value
Antenna array	ULA with spacing $\lambda/2 = 15$ cm
Number of array antennas	32
Number of UEs	2048
Carrier frequency	2.0 GHz
Bandwidth	312.5 kHz
SNR	0 dB

IV. Complexity Analysis Comparison

When we compare the complexity of our algorithm against the three algorithms used in [1], the most important advantage is that our algorithm does not require training or an abundant number of CSI to be able to reduce dimensionality efficiently. We can calculate the channel chart even for one UE data. This can make us calculate the channel chart sequentially (real-time) as the data is received, rather than store the data of 2048 UEs and use it all at once as in the case of PCA, SM, or AE, which consumes a very large amount of memory and complexity. The other advantage is the latency. Other algorithms need to collect the data of all UEs, which can take some time, and if the system is mobile, the geometry might have already changed by the time the channel chart is calculated. In our case, we can calculate each UE channel chart as we receive it, which makes our algorithm much more efficient. The computational complexity per UE in our algorithm is mostly due to the MUSIC algorithm complexity, which consists of calculating the covariance matrix, the decomposition of the eigenvalues, PMF of θ and the peak search, all of this can be approximated by $\mathcal{O}(M^2p + M^2N)$ as in [5] where in our case M is the number of antenna elements, p is the number of angles searched which in our case is 360 (0 to 180 in 1/2 degree steps), and N is the number of UEs which is 2048. This complexity is lower than that of PCA which is $\mathcal{O}(M^2N+M^3)$, SM which is $\mathcal{O}(N^2)$ or, AE is much more complex but hard to quantify exactly. The conclusion from this section is that our algorithm is much simpler, and has lower latency (suitable for the mobile environment), than other algorithms presented in [1]. In addition to being computationally simpler, our algorithm requires less memory.

V. Performance Comparison

We now compare the performance of our algorithm employing three cellular wireless communication channels that were used in [1]. In [1], these are called Vanilla LOS (LOS), Quadriga LOS (QLOS), and Quadriga NLOS (QNLOS) [6]. We will use the same nomenclature. For each channel we compare three scenarios discussed in Section III. The results are summarized in Table II for k-nearest neighbors equal to 102. As in [1] we use continuity (CT) and trustworthiness (TW) as performance measures. CT measures if neighbors in the original space are close in the representation space. TW measures how well the feature mapping avoids introducing

new neighbor relations that were absent in the original space. Let $\mathcal{V}_K(\mathbf{u}_i)$ be the K-neighborhood of point $\mathbf{u}_i$ in the original space. Also, let $\hat{r}(i,j)$ be the ranking of point $\mathbf{v}_j$ among the neighbors of point $\mathbf{v}_i$, ranked according to their similarity to $\mathbf{v}_i$. Then the point-wise continuity of the representation $\mathbf{v}_i$ of the point $\mathbf{u}_i$ is defined as

$$\mathrm{CT}_i(K) = 1 - \frac{2}{K(2N-3K-1)} \sum_{j \in \mathcal{V}_K(\mathbf{u}_i)} (\hat{r}(i,j) - K).$$

The (global) continuity a point set $\{\mathbf{u}_n\}_{n=1}^N$ and its representation $\{\mathbf{v}_n\}_{n=1}^N$ is $\mathrm{CT}(K) = \frac{1}{N}\sum_{i=1}^N \mathrm{CT}_i(K)$. Now, let $\mathcal{U}_K(\mathbf{v}_i)$ be the set of "false neighbors" that are in the K-neighborhood of $\mathbf{v}_i$, but not of $\mathbf{u}_i$ in the original space. Also, let $r(i,j)$ be the ranking of point $\mathbf{u}_i$ in the neighborhood of point $\mathbf{u}_i$, ranked according to their similarity to $\mathbf{u}_i$. The point-wise trustworthiness of the representation of point $\mathbf{u}_i$ is then

$$\mathrm{TW}_i(K) = 1 - \frac{2}{K(2N-3K-1)} \sum_{j \in \mathcal{U}_K(\mathbf{v}_i)} (r(i,j) - K).$$

The (global) trustworthiness between a point set $\{\mathbf{u}_n\}_{n=1}^N$ and its representation $\{\mathbf{v}_n\}_{n=1}^N$ is $\mathrm{TW}(K) = \frac{1}{N}\sum_{i=1}^N \mathrm{TW}_i(K)$. Both point-wise and global CT and TW are between 0 and 1, with larger values being better [1]. One can see from Table II that LR and ISQ substantially outperform the techniques in [1], namely PCA, SM, and AE. Comparing our two techniques LR and ISQ, we find that LR (supervised) is only slightly better than ISQ (less than 0.5% in the case of LOS and less than 0.1% in the case of QLOS and QNLOS). This is negligible compared to the overhead of using GPS and relaying this information to the BS. We note that the reason we look at less than 100% area scale in Table II and in the sequel is the possibility of getting sufficiently good performance with smaller distances.

Fig. 2 and Fig. 3 present the channel charts. Fig. 2 presents the case for 100% of the original dimensions and Fig. 3 presents the case for 25% of the original dimensions. In both figures, columns 1 through 4 correspond to PCA, SM, AE, and MUSIC (ISQ) algorithms. Therefore, one should compare the fourth column with the other three columns on a row by row basis. The considered system geometry is given in [1, Fig. 1(a)]. The goal of the channel chart is to employ CSI and then derive a chart which preserves the distances in the system geometry. It can be seen from Fig. 2 and Fig. 3 that our algorithm does a significantly better job than PCA, SM, and AE in that regard. In particular, the letters VIP present in [1, Fig. 1(a)] can be seen to be much more preserved with our algorithm.

Fig. 4 shows the correlation of the estimated ρ with that of the real ρ. Both methods correlate with real ρ almost linearly.

Fig. 5–Fig. 7 plot TW and CT values for PCA, SM, AE, ISQ, and LR algorithms. Fig. 5 is for 100% of the original dimensions, Fig 6 is for 50% of the original dimensions, Fig. 7 is for 25% of the original dimensions. The horizontal axis is for k-nearest neighbors and the number varies between 0 and 100. The first through third columns are for the LOS, QLOS, and QNLOS channels, respectively. For the LOS channel, ISQ and LR substantially outperform the other three algorithms with negligible difference between the two. This is especially true for TW. For the QLOS and QNLOS channels, and for CT, LR and ISQ beat the other three, with LR being better than ISQ. For the same channels, for TW, LR and ISQ have close performance yet they are substantially better than the remaining three.

TABLE II
PERFORMANCE COMPARISON FOR TW AND CT AT k-nearest = 102

Measure/Area	Scale/Channel		PCA	SM	AE	LR	ISQ
TW	1	LOS	0.8565	0.7986	0.8303	**0.9929**	0.9887
		QLOS	0.8447	0.8359	0.8655	0.9119	**0.9125**
		QNLOS	0.8500	0.8459	0.8533	0.9090	**0.9095**
	1/2	LOS	0.8596	0.8200	0.8349	**0.9932**	0.9889
		QLOS	0.8579	0.8343	0.8510	**0.9056**	0.9050
		QNLOS	0.8682	0.8661	0.8747	0.9376	**0.9386**
	1/4	LOS	0.8570	0.7616	0.8118	**0.9930**	0.9886
		QLOS	0.8630	0.8724	0.8667	**0.9494**	0.9489
		QNLOS	0.8842	0.8885	0.8997	0.9544	**0.9545**
CT	1	LOS	0.9270	0.8749	0.8921	**0.9967**	0.9941
		QLOS	0.9219	0.9007	0.8718	**0.9445**	0.9360
		QNLOS	0.9250	0.9242	0.9158	**0.9339**	0.9315
	1/2	LOS	0.9279	0.9004	0.8915	**0.9969**	0.9941
		QLOS	0.9343	0.9047	0.9296	**0.9604**	0.9504
		QNLOS	0.9331	0.9296	0.9210	0.9516	**0.9529**
	1/4	LOS	0.9271	0.8432	0.8616	**0.9968**	0.9940
		QLOS	0.9348	0.9390	0.9297	**0.9699**	0.9688
		QNLOS	0.9360	0.9377	0.9325	0.9593	**0.9602**

VI. CONCLUSION AND FUTURE WORK

The algorithm presented in this letter significantly outperforms the three algorithms in the seminal paper [1], both in performance and computational complexity. The improvement can reach 23% in some cases. The most important advantage of our algorithm over other dimensionality reduction algorithms is that we can calculate each UE data independently as it comes, so it is much faster and simpler. We assumed a 2-D environment and we used static channels and single sub carrier CSI as in [1].

REFERENCES

[1] C. Studer, S. Medjkouh, E. Gonultas, T. Goldstein, and O. Tirkkonen, "Channel charting: Locating users within the radio environment using channel state information," *IEEE Access*, vol. 6, pp. 47 682–47 698, 2018.

[2] L. Le Magoarou, "Efficient channel charting via phase-insensitive distance computation," *IEEE Wireless Communications Letters*, vol. 10, no. 12, pp. 2634–2638, 2021.

[3] Q. Zhang and W. Saad, "Semi-supervised learning for channel charting-aided IoT localization in millimeter wave networks," in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–6.

[4] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[5] C. Stoeckle, J. Munir, A. Mezghani, and J. A. Nossek, "DoA estimation performance and computational complexity of subspace- and compressed sensing-based methods," in *WSA 2015; 19th International ITG Workshop on Smart Antennas*, 2015, pp. 1–6.

[6] S. Jaeckel, L. Raschkowski, K. Borner, and L. Thiele, "Quadriga: A 3-D multi-cell channel model with time evolution for enabling virtual field trials," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 6, p. 3242–3256, 2014.

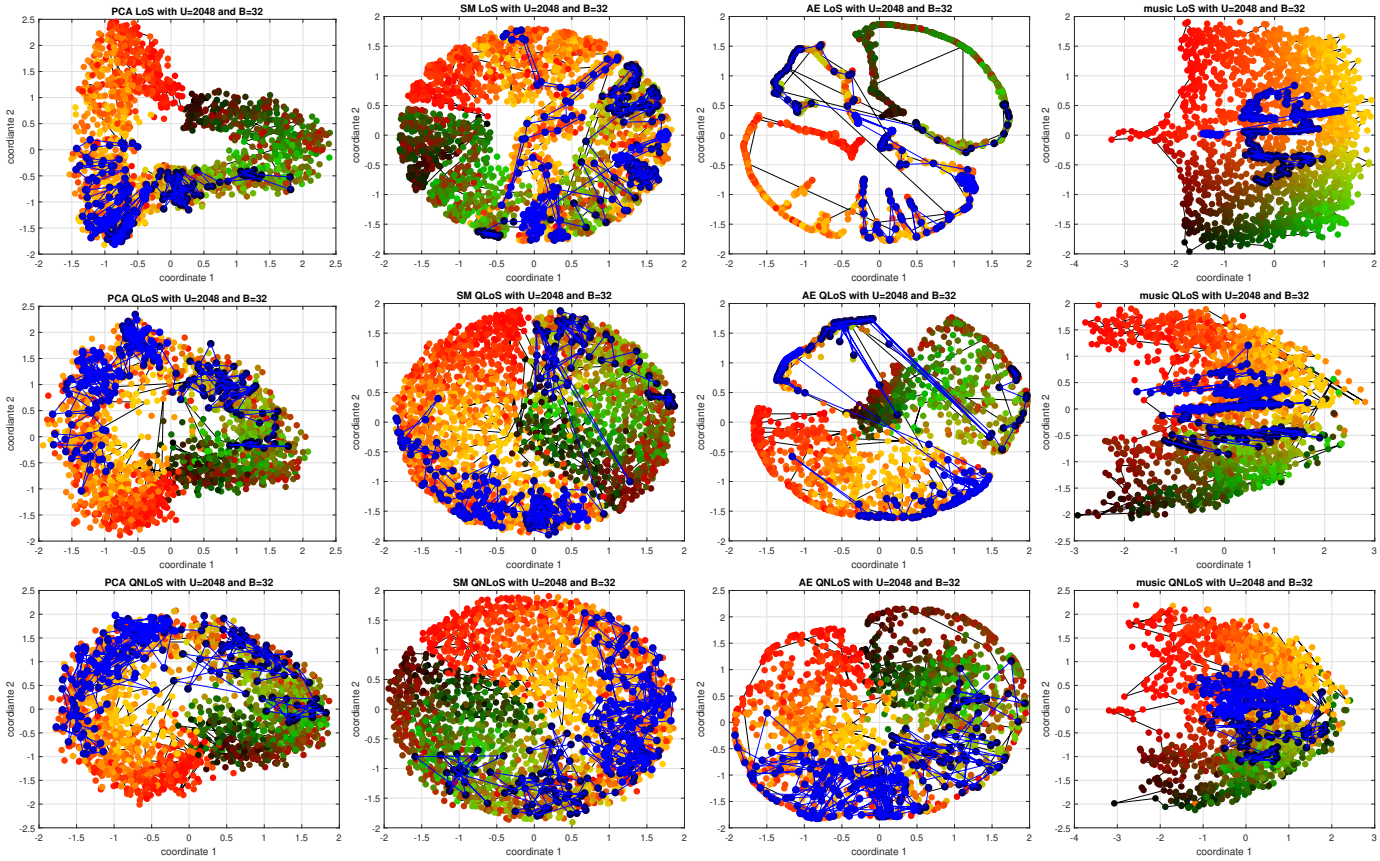

Fig. 2. Original dimensions.

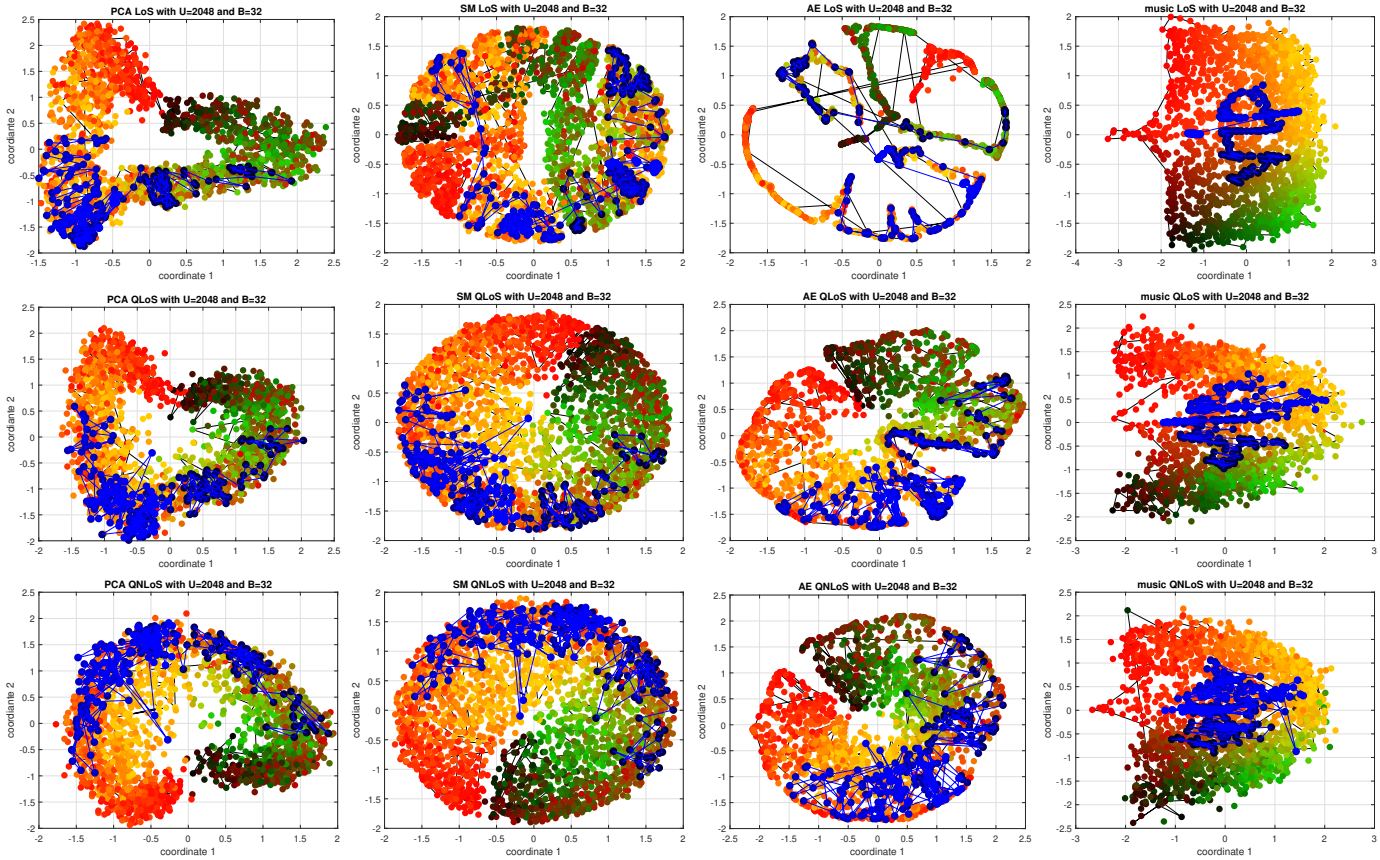

Fig. 3. 25% of original dimensions.

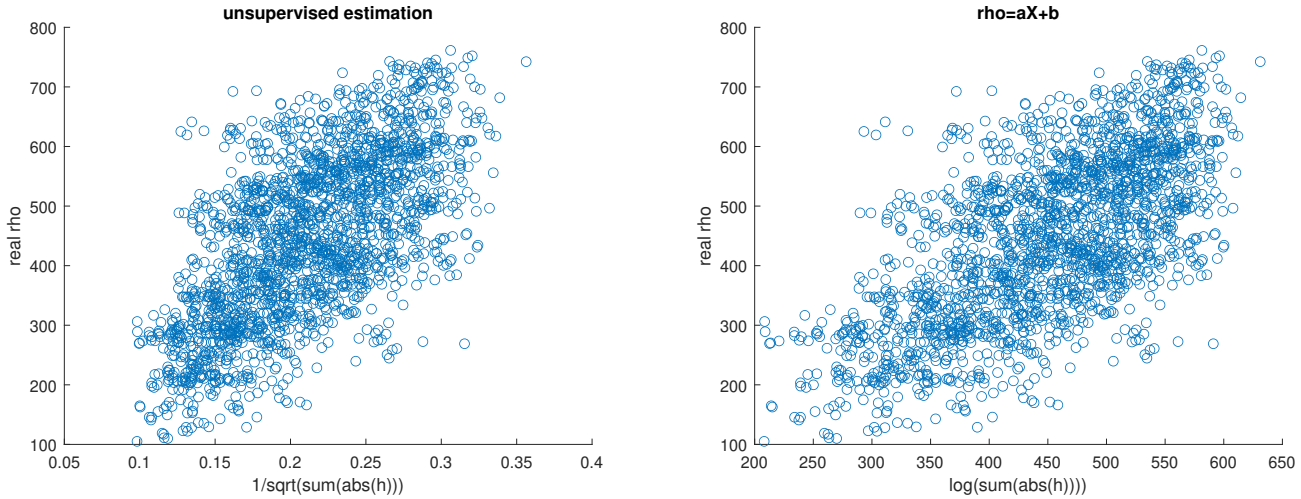

Fig. 4. Correlation of real ρ vs estimated ρ.

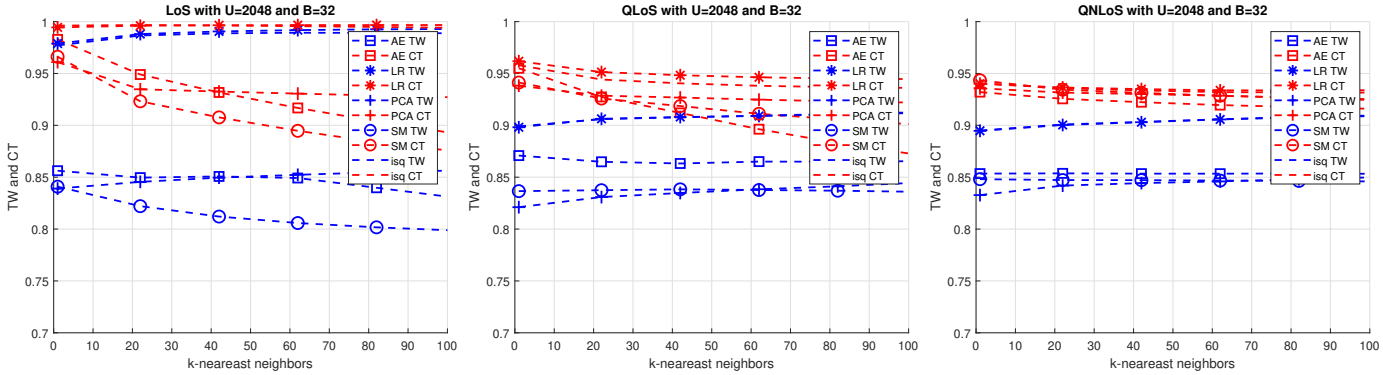

Fig. 5. TW and CT performance of PCA, SM, AE, ISQ, and LR, original dimensions.

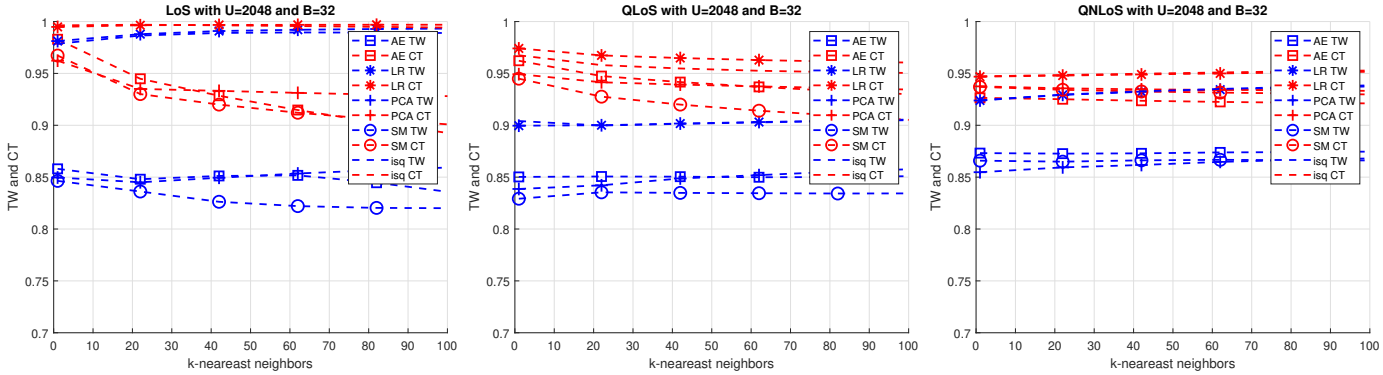

Fig. 6. TW and CT performance of PCA, SM, AE, ISQ, and LR, 50% of original dimensions.

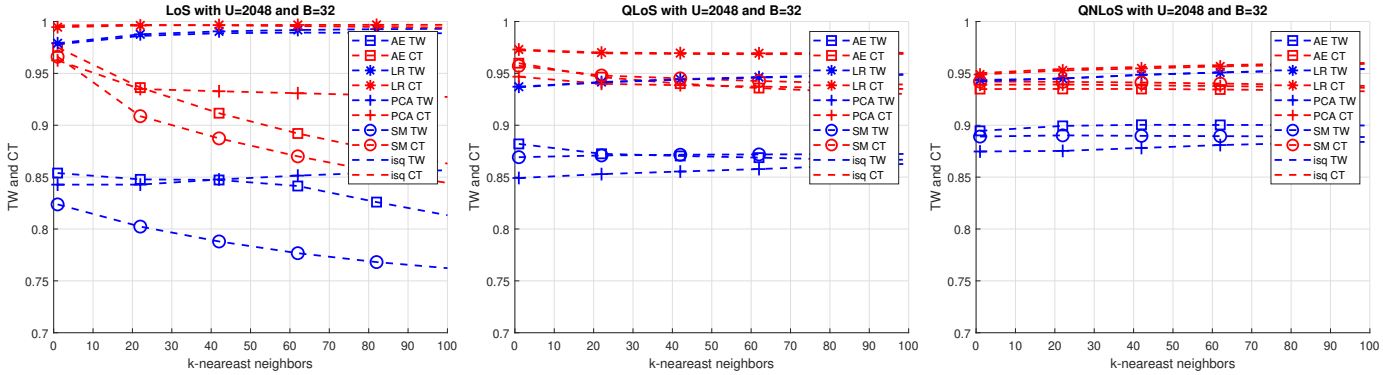

Fig. 7. TW and CT performance of PCA, SM, AE, ISQ, and LR, 25% of original dimensions.