CHAPTER SIX

MAXIMUM ENTROPY MODELLING

6.1 INTRODUCTION

It is well established that multiple-input-multiple-output (MIMO) wireless systems have the potential to meet demands for higher data rates, improved spectral efficiency and better quality of service in wireless systems [15]. Correct assessment of MIMO systems depends on accurate characterization of the propagation channel, which has resulted in numerous measurement campaigns [89, 99–101] as well as modelling strategies [17, 19].

One of the most popular MIMO modelling strategies to date is the Kronecker model [18], which allows the full joint transmit/receive covariance matrix to be synthesized from only the separate transmit and receive covariance matrices. Although the model is preferred due to its simplicity, application of the model to measured channels indicates marginal discrepancies in the modelled capacity and more substantial error in the modelled joint spatial spectra (eigenvalues) [97]. Improvements to the Kronecker model have been proposed by combining the separate transmit/receive covariance information with partial information from the full covariance [102].

A modelling approach based on information theoretic considerations has appeared [103], which applies the principle of maximum entropy (ME) to MIMO channels. Given a set of known channel parameters, the ME estimate of any unknown parameters is that which imposes the least structure on the channel model. Thus, by not imposing "artificial structure" on the channel, the ME principle should allow for more consistent (and possibly more accurate)

modelling, based on only the channel knowledge at hand.

The purpose of the present work is to apply the principle of ME to the problem of deriving the full joint covariance, based only on knowledge of the separate transmit/receive covariances. The main goal is to see if ME offers any modelling improvement by avoiding the artificial structure apparently present in the Kronecker model [97].

6.2 MODEL DESCRIPTION

Suppose the respective transmit and receive covariances $\mathbf{R}_T$ and $\mathbf{R}_R$ are known, respectively, so then the statistics of the channel matrices are constrained by

$$\mathrm{E}\left\{\mathbf{H}\mathbf{H}^H\right\} = \mathbf{R}_R \quad \mathrm{E}\left\{\sum_k H_{ik}H_{jk}^*\right\} = R_{R,ij} \tag{6.1}$$

$$\mathrm{E}\left\{\mathbf{H}^H\mathbf{H}\right\} = \mathbf{R}_T^T \quad \mathrm{E}\left\{\sum_k H_{ki}H_{kj}^*\right\} = R_{T,ij} \tag{6.2}$$

$$p(\mathbf{H}) \geq 0 \tag{6.3}$$

$$\int p(\mathbf{H})d\mathbf{H} = 1, \tag{6.4}$$

where $\{\cdot\}^H$ is complex conjugate transpose, $p(\mathbf{H})$ is the joint probability density function (PDF) of the elements of the channel matrix $\mathbf{H}$, and $\mathrm{E}\{\cdot\}$ is expectation. Assuming we want to maximize the entropy or

$$\int p(\mathbf{H}) \log p(\mathbf{H})d\mathbf{H} \tag{6.5}$$

with respect to the above constraints, one can write the Lagrangian and set the derivative equal to zero to get the form of the PDF

$$\begin{aligned} & p(\mathbf{H}) \\ &= \exp(\lambda_0 + \sum_{ij}\mu_{R,ij}\sum_k H_{ik}H_{jk}^* + \sum_{ij}\mu_{T,ij}\sum_k H_{ki}H_{kj}^*) \\ &= c_0 \exp(\sum_{ijk\ell}\mu_{R,ik}H_{ij}H_{k\ell}^*\delta_{j\ell} + \sum_{ijk\ell}\mu_{T,j\ell}H_{ij}H_{k\ell}^*\delta_{ik}), \end{aligned} \tag{6.6}$$

where $c_0 = \exp(\lambda_0)$. Combining the two equations leads to

$$p(\mathbf{H}) = c_0 \exp\left[\sum_{ijkl} H_{ij}H_{k\ell}^*(\underbrace{\mu_{R,ik}\delta_{j\ell} + \mu_{T,j\ell}\delta_{ik}}_{-R_{ij,k\ell}^{-1}})\right], \tag{6.7}$$

which is the form of the standard multivariate complex normal PDF with a covariance matrix indexed by the stacked receive and transmit indices. This can be expressed with the Kronecker product as

$$\mathbf{R} = -(\underbrace{\mathbf{I}_T \otimes \boldsymbol{\mu}_R + \boldsymbol{\mu}_T \otimes \mathbf{I}_R}_{\mathbf{A}})^{-1} \tag{6.8}$$

One observes that (6.8) is different from the Kronecker model [18] which has the simpler form

$$\mathbf{R} = \mathbf{R}_T \otimes \mathbf{R}_R. \tag{6.9}$$

Taking the eigenvalue decomposition (EVD) of $\boldsymbol{\mu}_T$ and $\boldsymbol{\mu}_R$, ie.,

$$\boldsymbol{\mu}_T = \boldsymbol{\xi}_T \boldsymbol{\Lambda}_T \boldsymbol{\xi}_T^H, \quad \boldsymbol{\mu}_R = \boldsymbol{\xi}_R \boldsymbol{\Lambda}_R \boldsymbol{\xi}_R^H, \tag{6.10}$$

allows (6.8) to be expanded as

$$\mathbf{R}^{-1} = -\underbrace{(\boldsymbol{\xi}_T \otimes \boldsymbol{\xi}_R)}_{\boldsymbol{\xi}} \underbrace{(\mathbf{I}_T \otimes \boldsymbol{\Lambda}_R + \boldsymbol{\Lambda}_T \otimes \mathbf{I}_R)}_{\boldsymbol{\Lambda}} \underbrace{(\boldsymbol{\xi}_T \otimes \boldsymbol{\xi}_R)^H}_{\boldsymbol{\xi}^H}, \tag{6.11}$$

so that $\mathbf{R} = -\boldsymbol{\xi}\boldsymbol{\Lambda}^{-1}\boldsymbol{\xi}^H$. The matrix $\boldsymbol{\Lambda}$ involves Kronecker products and sums of diagonal matrices, which will also be diagonal, or

$$\begin{aligned} \Lambda_{ij,k\ell} &= \delta_{j\ell}\delta_{ik}\lambda_{R,i} + \delta_{j\ell}\delta_{ik}\lambda_{T,j} \\ &= \delta_{j\ell}\delta_{ik}\left[\lambda_{R,i} + \lambda_{T,j}\right] \qquad (6.12) \\ \Lambda^{-1}_{ij,k\ell} &= \delta_{j\ell}\delta_{ik}\frac{1}{\lambda_{R,i} + \lambda_{T,j}} \qquad (6.13) \end{aligned}$$

Therefore one has the form of the full covariance matrix. The eigenvectors are just the Kronecker product of the separate transmit and receive eigenvectors. The eigenvalues, however, must be found by substituting $\mathbf{R}$ back into the original constraints. Although the constraints (6.1) and (6.2) depend on expectations of $\mathbf{H}$, they can be re-written in terms of the full covariance as

$$R_{R,ik} = \sum_j R_{ij,kj}, \quad R_{T,j\ell} = \sum_i R_{ij,i\ell} \tag{6.14}$$

The covariance matrix from (6.11) can be written in component form as

$$R_{ij,k\ell} = \left\{-\mathbf{A}^{-1}\right\}_{ij,k\ell} = -\sum_{mn} \xi_{ij,mn}\Lambda^{-1}_{mn,mn}\xi^*_{k\ell,mn} \tag{6.15}$$

Substituting (6.15) into (6.14) results in

$$\{-\mathbf{R}_T\}_{jl} = \sum_{mn}\sum_{i} \xi_{ij,mn}\Lambda^{-1}_{mn,mn}\xi^*_{i\ell,mn} \tag{6.16}$$

$$= \sum_{mn} \xi_{T,jn}\xi^*_{T,\ell n}\Lambda^{-1}_{mn,mn} \underbrace{\sum_{i} \xi_{R,im}\xi^*_{R,im}}_{=1} \tag{6.17}$$

$$= \sum_{n} \xi_{T,jn}\xi^*_{T,\ell n} \underbrace{\sum_{m} \Lambda^{-1}_{mn,mn}}_{-D_{T,nn}}, \tag{6.18}$$

Similarly, one can expand the receive covariance constraint as

$$\{-\mathbf{R}_R\}_{ik} = \sum_{m} \xi_{R,im}\xi^*_{R,km} \underbrace{\sum_{n} \Lambda^{-1}_{mn,mn}}_{-D_{R,mm}} \tag{6.19}$$

Equations (6.18) and (6.19) can be recognized as just the eigenvalue decompositions of the transmit and receive covariances

$$\mathbf{R}_T = \boldsymbol{\xi}_T\mathbf{D}_T\boldsymbol{\xi}_T^H \tag{6.20}$$

$$\mathbf{R}_R = \boldsymbol{\xi}_R\mathbf{D}_R\boldsymbol{\xi}_R^H. \tag{6.21}$$

To solve for the full covariance, the system of equations

$$D_{T,nn} = d_{T,n} = -\sum_{m} \frac{1}{\lambda_{R,m} + \lambda_{T,n}} \tag{6.22}$$

$$D_{R,mm} = d_{R,m} = -\sum_{n} \frac{1}{\lambda_{R,m} + \lambda_{T,n}} \tag{6.23}$$

must be solved, but a direct solution can be problematic since one would have to find the roots of multivariate large order polynomials. An indirect approach is possible by noticing that since $\mathbf{H}$ is a Gaussian process, maximum entropy maximizes $det(\mathbf{R})$. Since the eigenvectors are already known, one only needs to find $\boldsymbol{\Lambda}$, such that $\det\left(-\boldsymbol{\Lambda}^{-1}\right)$ is maximized. Letting

$$f_{mn} = -\Lambda^{-1}_{mn,mn}, \tag{6.24}$$

one must find the maximum of $\det(\mathbf{R}) = \prod_{ij} f_{ij}$, subject to the constraints

$$d_{T,j} = \sum_i f_{ij} \tag{6.25}$$

$$d_{R,i} = \sum_j f_{ij} \tag{6.26}$$

$$f_{ij} \geq 0 \tag{6.27}$$

Since the constraints are linear and $-\prod_{ij} f_{ij}$ is convex, there is a simple convex optimization problem that can be solved with conventional techniques. In this research, the solution is obtained by finding an admissible initial guess for the f_{ij} using linear programming and then moving to the optimum with gradient descent method.

6.3 DATA PROCESSING

The data is processed as follows: At each location, 20 channel snapshots were recorded with 200 ms between snapshots. Since negligible channel variation was observed for each stationary measurement, only a single snapshot from each location was considered. Here, a channel snapshot is defined as $H_{ij}^{(k)}$, where k is a frequency bin index, and i and j are the receive and transmit antenna indices respectively. To remove the effect of path loss in the computations, channel matrices were normalized to have average unit SISO gain [95].

Previous channel modelling efforts have defined the double-directional channel [17] in terms of paired discrete plane-wave departures and arrivals at the TX and RX. In indoor environments, where multipath scattering is severe, extracting individual plane-wave arrivals can be very difficult. It was therefore decided to define the double-directional response in terms of spatial power spectra, obtained with joint TX/RX Bartlett beamformers. The joint Bartlett beamformer [42] is given by

$$P(\nu_T, \nu_R) = \frac{\mathbf{a}(\nu_T, \nu_R)^H \hat{\mathbf{R}} \mathbf{a}(\nu_T, \nu_R)}{\mathbf{a}(\nu_T, \nu_R)^H \mathbf{a}(\nu_T, \nu_R)} \tag{6.28}$$

where ν_T and ν_R are azimuth angles at the TX and RX, and $\hat{\mathbf{R}}$ is the sample covariance matrix. The joint steering vector $\mathbf{a}(\nu_T, \nu_R)$ is defined as

$$\mathbf{a}(\nu_T, \nu_R) = \mathbf{a}_T(\nu_T) \otimes \mathbf{a}_R(\nu_R), \tag{6.29}$$

where $\mathbf{a}_{\{T,R\}}$ is the usual separate array steering vectors for the TX and RX, and $\otimes$ is the Kronecker product. The sample covariance matrix is computed as

$$\hat{\mathbf{R}} = \frac{1}{K}\sum_{k} \mathbf{h}^{(k)}\mathbf{h}^{(k)H}, \tag{6.30}$$

where K is the total number of frequency bins, $h^{(k)} = \mathrm{Vec}\left\{H^{(k)}\right\}$, and the vector operation $\mathrm{Vec}\left\{\cdot\right\}$ stacks a matrix into a vector.

The similarity of the joint spectra at 2.4 GHz is evaluated by computing the correlation coefficient between spectra A and B as

$$\rho_{A,B} = \frac{\sum_{j=0}^{N}\sum_{i=0}^{N}(P_{A,ij}-\bar{P}_A)(P_{B,ij}-\bar{P}_B)}{\sqrt{\left[\sum_{i=0}^{N}\sum_{j=0}^{N}(P_{A,ij}-\bar{P}_A)^2\right]\left[\sum_{i=0}^{N}\sum_{j=0}^{N}(P_{B,ij}-\bar{P}_B)^2\right]}}, \tag{6.31}$$

where N is the number of discretization points, $P_{ij} = P(\nu_{T,i}, \nu_{R,j})$, A and B represent FC for full covariance extracted from the data, KM for the Kronecker model, and ME for maximum entropy, $\nu_{T,i} = \nu_{R,i} = 2\pi i/N$, and $\bar{P} = \left(\frac{1}{N^2}\right)\sum_i\sum_j P_{ij}$.

6.4 RESULTS

Figures 6.1and 6.2 display the significant eigenvalues for the full covariance (FC), maximum entropy (ME), and the Kronecker model (KM) cases, at Locations 3 and 7 at the carrier frequency of 2.4 GHz for the indoor environment in Figure 4.21. Figures 6.3 and 6.4 display the significant eigenvalues for the same physical locations, but at the carrier frequency of 5.2 GHz. All these cases indicate that the maximum entropy approach only provides a marginal improvement in the eigenvalues (approximately 5%) compared to the Kronecker model. The other nine locations were very similar, except that only the number of dominant eigenvalues varied from around a minimum of two to a maximum of four.

Figure 6.5 plots the spatial spectra for the full covariance (top) and maximum entropy (bottom). The spatial spectrum for the Kronecker model (not plotted) is visibly indistinguishable from that for the maximum entropy. Tables 6.1 and 6.2 give the correlation coefficient $\rho_{FC,ME}$ and $\rho_{FC,KM}$ in columns 2 and 3 for the center frequencies of 2.4 GHz and 5.2 GHz respectively.

TABLE 6.1: Correlation coefficient of spatial power spectra at 2.4 GHz

Locations	Bartlett beamforming at 2.4 GHz	
	$\rho_{FC,ME}$	$\rho_{FC,KM}$
1	0.9387	0.9377
2	0.9258	0.9088
3	0.9470	0.9264
4	0.9927	0.9887
5	0.9716	0.9615
6	0.9818	0.9798
7	0.8850	0.8819
8	0.9872	0.9845
9	0.9028	0.8966
10	0.9007	0.8985
11	0.9261	0.9229

TABLE 6.2: Correlation coefficient of spatial power spectra at 5.2 GHz

Locations	Bartlett beamforming at 5.2 GHz	
	$\rho_{FC,ME}$	$\rho_{FC,KM}$
1	0.9701	0.9764
2	0.9675	0.9664
3	0.9109	0.9064
4	0.9945	0.9907
5	0.9717	0.9680
6	0.9694	0.9691
7	0.8411	0.8403
8	0.9936	0.9921
9	0.9637	0.9584
10	0.9363	0.9259
11	0.8339	0.8372

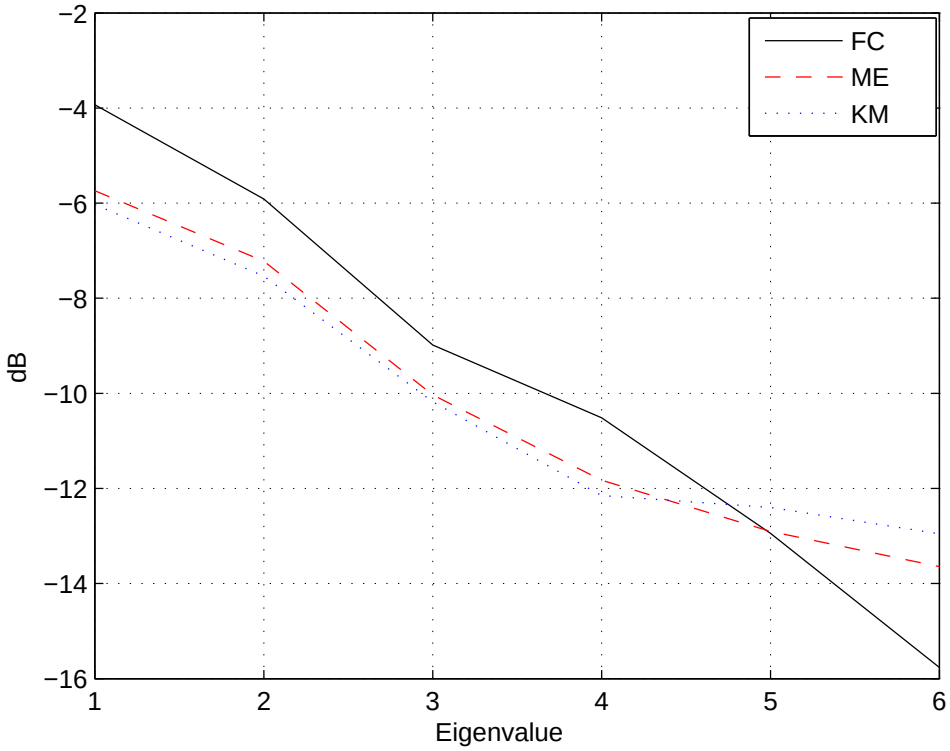

FIGURE 6.1: Dominant singular eigenvalues for Location 3 at 2.4 GHz carrier frequency

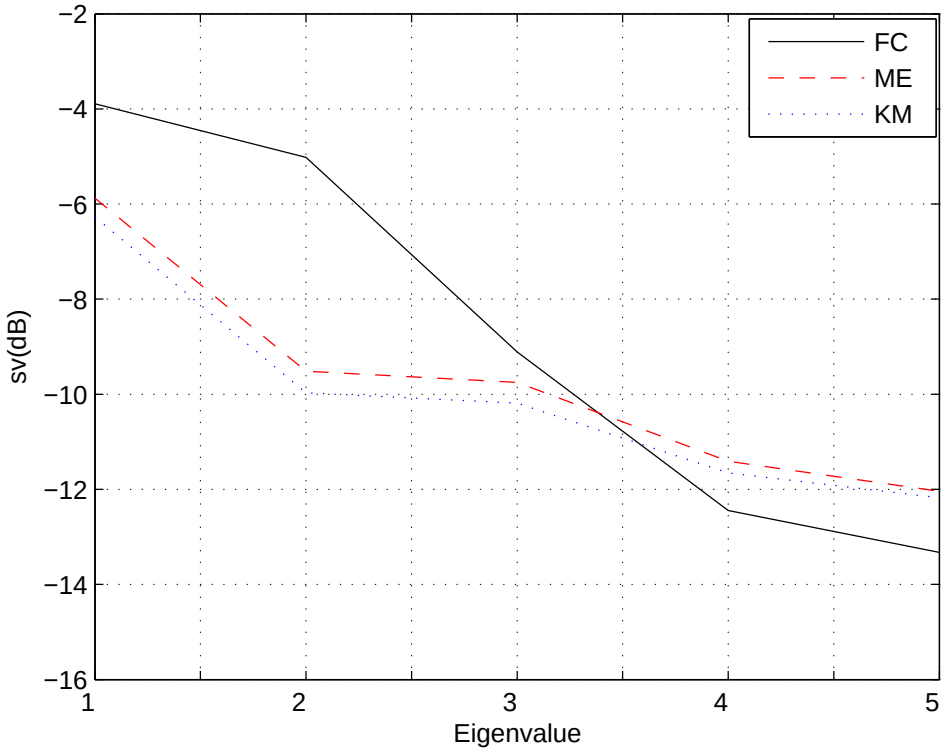

FIGURE 6.2: Dominant singular eigenvalues for Location 7 at 2.4 GHz carrier frequency

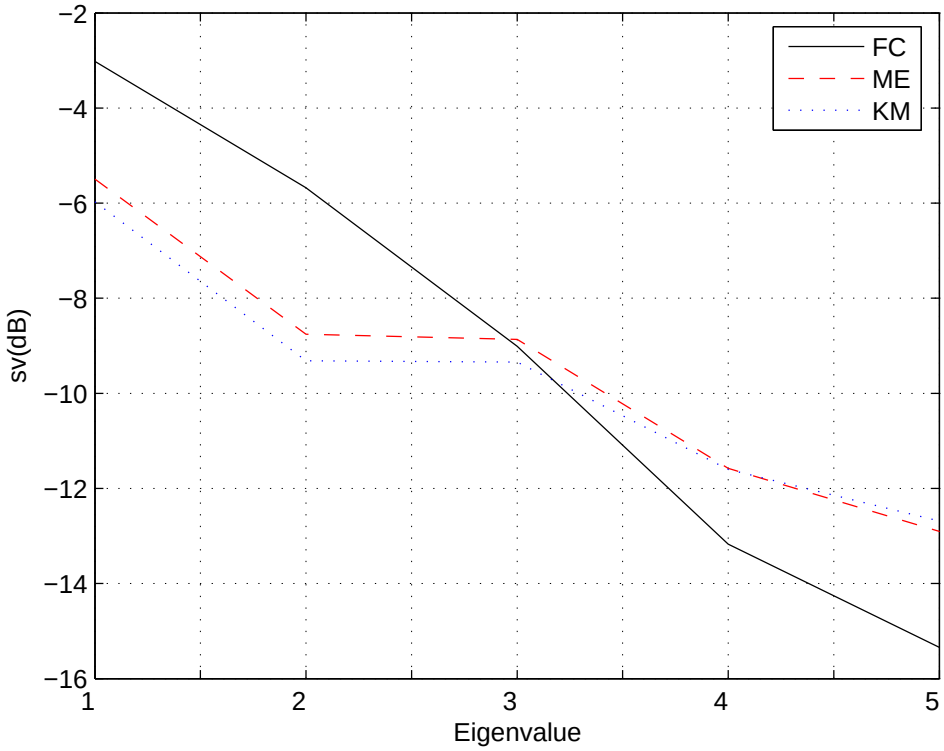

FIGURE 6.3: Dominant singular eigenvalues for Location 3 at 5.2 GHz carrier frequency

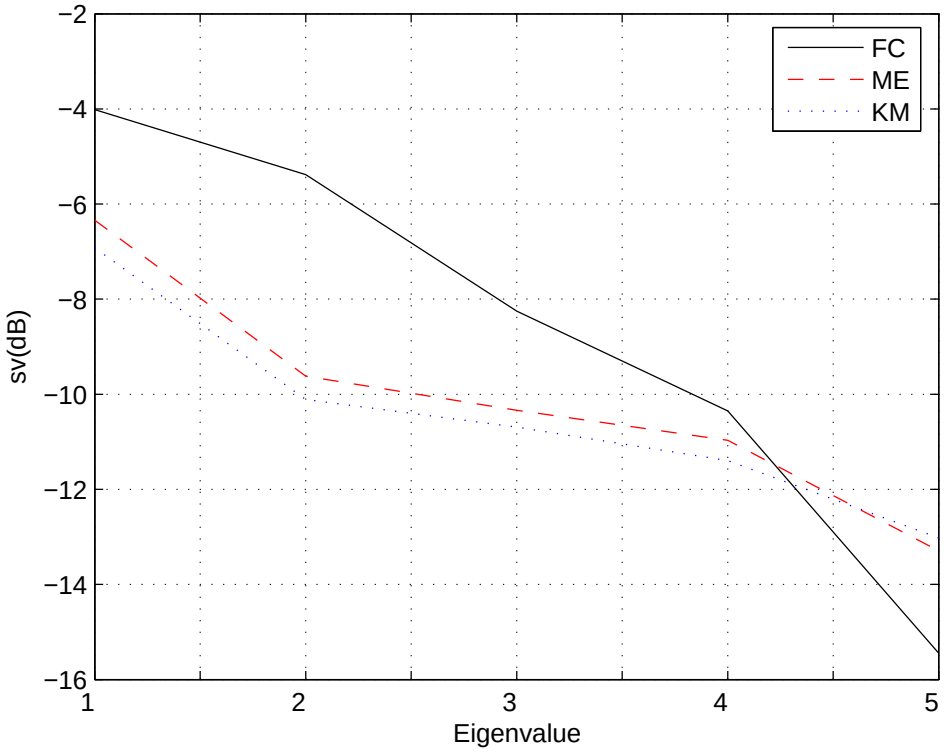

FIGURE 6.4: Dominant singular eigenvalues for Location 7 at 5.2 GHz carrier frequency

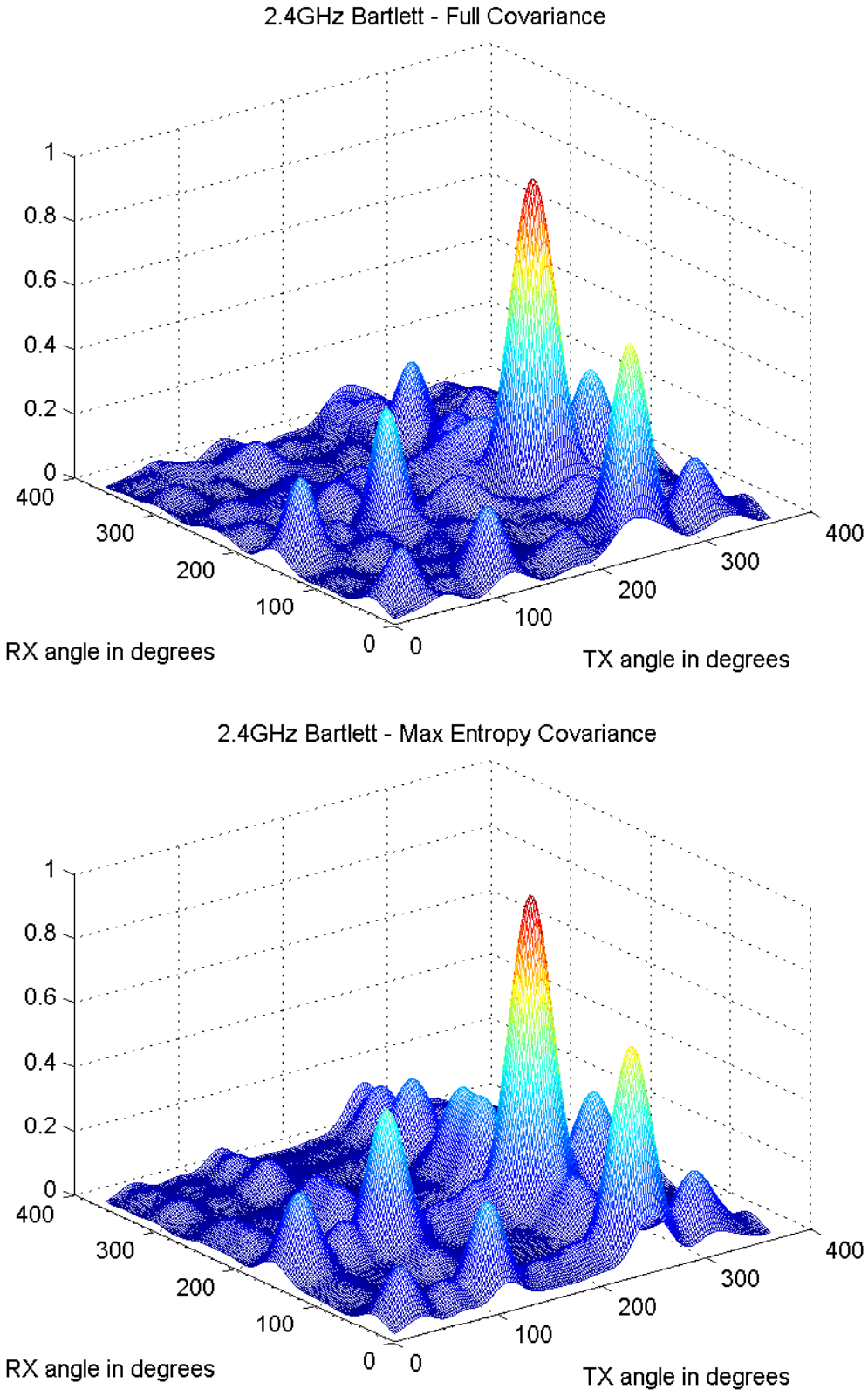

FIGURE 6.5: Spatial power spectra for FC and ME at 2.4 GHz at Location 3

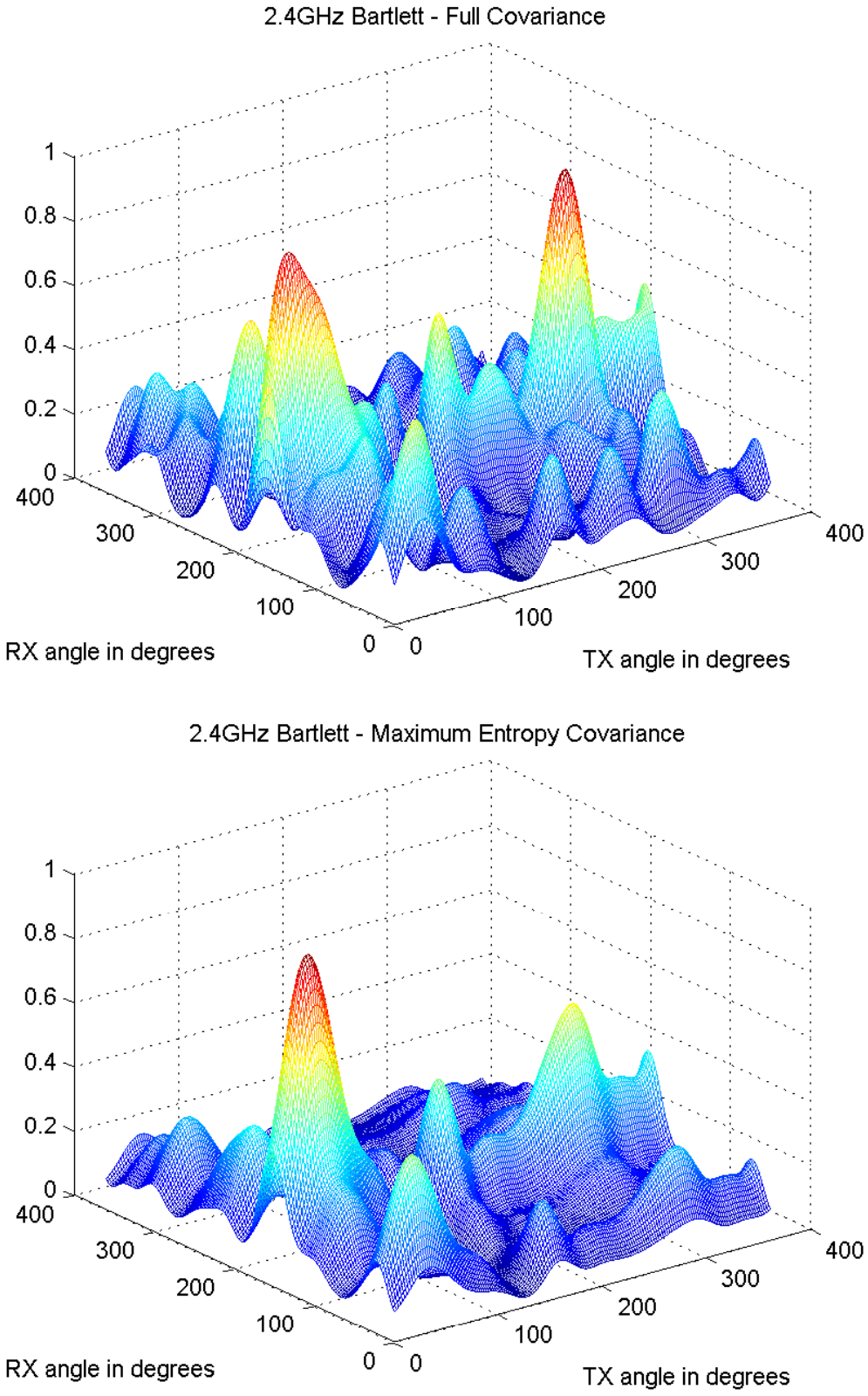

FIGURE 6.6: Spatial power spectra for FC and ME at 2.4 GHz at Location 7

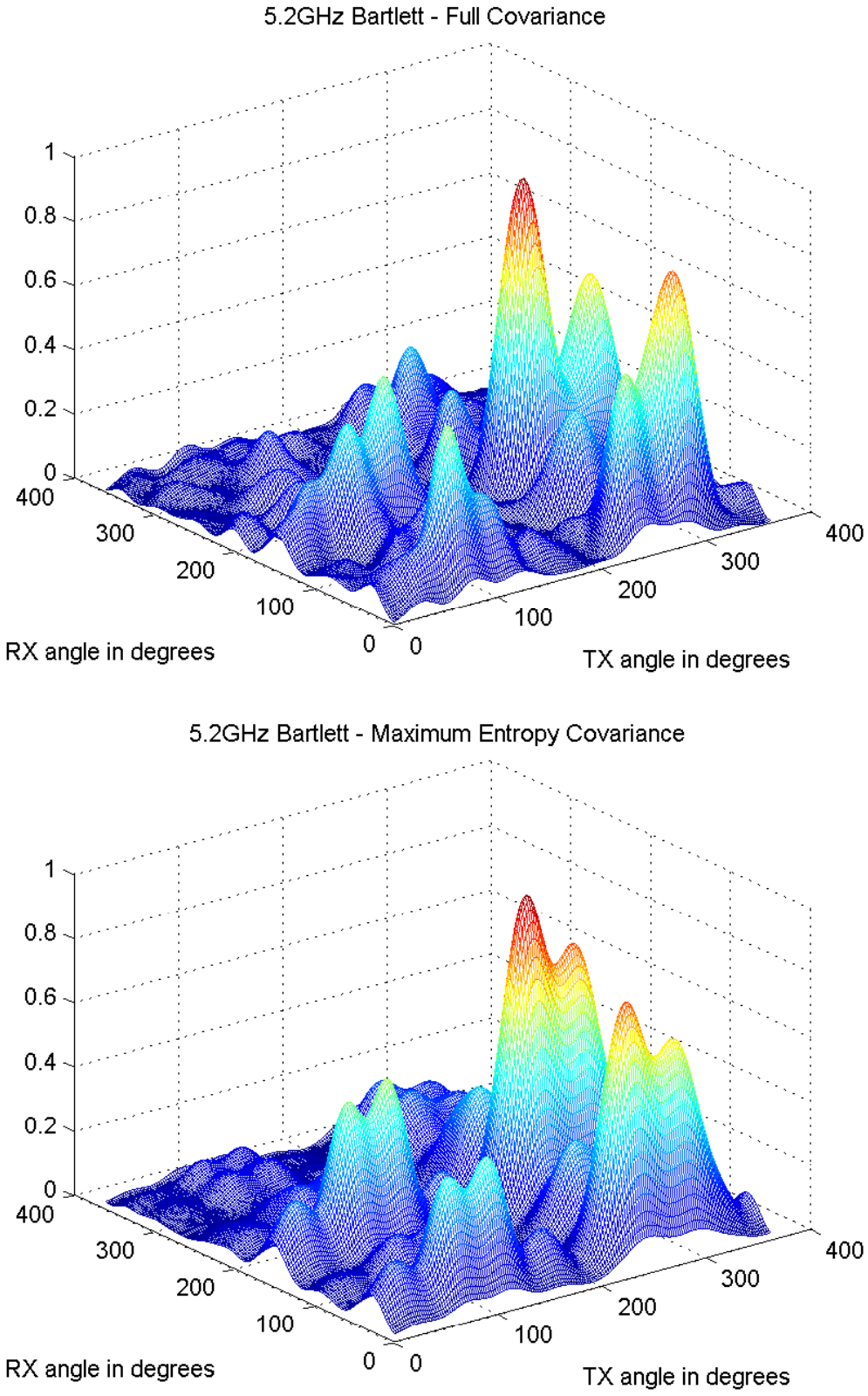

FIGURE 6.7: Spatial power spectra for FC and ME at 5.2 GHz at Location 3

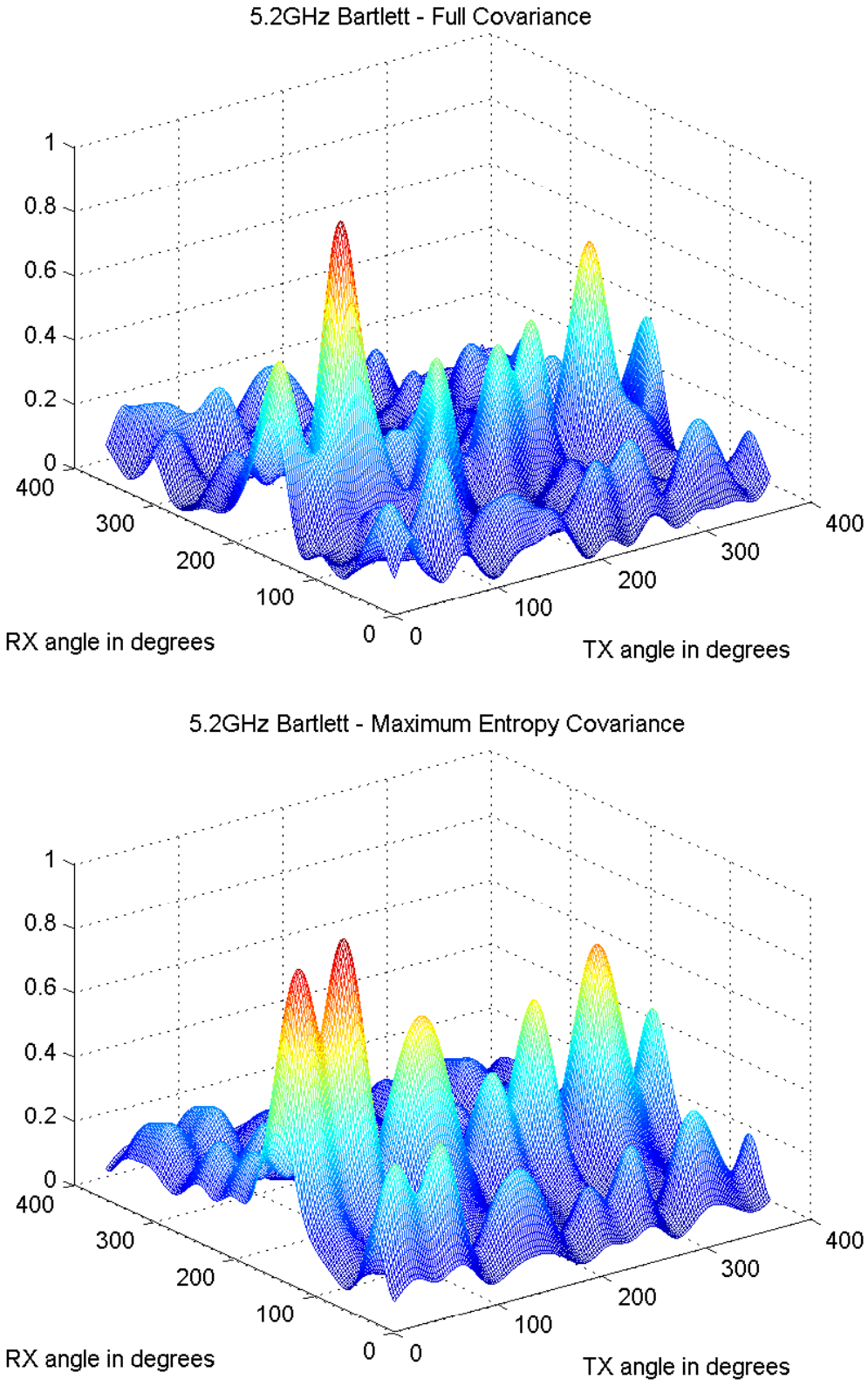

FIGURE 6.8: Spatial power spectra for FC and ME at 5.2 GHz at location 7

As with the eigenvalues, the maximum entropy approach only yields a marginal improvement, yet is distinct from the Kronecker model.

6.5 CONCLUSION

A maximum entropy approach has been presented for obtaining the full channel covariance when only the separate transmit and receive covariances are known. The method indicates that the eigenvectors of the full covariance are the Kronecker product of the separate transmit and receive eigenvectors, which are identical to that of the Kronecker model. The eigenvalues of the full covariance, however, do not strictly match those of the Kronecker model. Application of the new method to measured indoor MIMO channels at 2.4 GHz indicated that the Kronecker model and maximum entropy method gave numerical results that are fairly close to each other. Thus, even though the Kronecker model imposes "structure" on the full covariance, the modelled channel matrices are nearly as unstructured as can be and they almost attain the maximum possible entropy. This result suggests that the inaccuracy of the Kronecker model represents a fundamental limit, which resulted from having no joint transmit/receive channel information, rather than resulting from the incorrect application of knowledge of the separate transmit and receive covariances.

CHAPTER SEVEN
CONCLUSION

7.1 SUMMARY

From the work described herein, one can conclude:

- Geometric modelling

  Chapter 3 chartered the development of a geometric model for the MIMO indoor wireless channel with isotropic scattering PDF at the TX and the von Mises scattering PDF applied to one RX end, while a uniform scattering PDF was applied at the other. The effects of antenna element spacing, angular orientation of antenna array and number of antenna elements employed were shown. It was also found that the number of antenna elements had the greatest impact on the capacity when compared to the other parameters. Although geometric modelling is a useful tool and gives the wireless engineer much insight into the performance of the MIMO channel, it was also shown that, based on the measurement campaign, the results from the geometrical modelling technique did not prove to be able to accurately represent the real world channel in this case.

- Wideband measurement system

  Chapter 4 showed a successful development and implementation of a low cost wideband 8 x 8 MIMO measurement system capable of 80 MHz of instantaneous excitation bandwidth operating across the 2-8 GHz range. The system shows innovation in its implementation of the sub-components and data was obtained for 11 different indoor locations both at 2.4 GHz and 5.2 GHz, using the ULA and UCA antenna configurations

in co-located positions. This measurement system is largely based on off-the-shelf components, making it affordable for empirical research to be undertaken.

- Metrics in data analysis

  In Chapter 5, the data was analyzed in terms of eigenvalue CDFs, capacities and spatial correlation for the specific antenna array configuration at the two carrier frequencies. The concept of *frequency scaling* in MIMO systems was introduced and two new different models were developed where it was shown that a linear relationship existed both for the capacity frequency scaling and spatial correlation frequency scaling. When the fixtures and fittings becomes electrically comparable to the wavelength of frequency in use, this *frequency scaling* approach could not hold. Hence, one should use these models taking careful cognisance of the center frequency in use.

  It was also shown that the capacity at 2.4 GHz was greater than that at 5.2 GHz for the UCA, which could be attributed to more scattering and lesser path loss. Furthermore one can observe that the capacity for the ULA was more clustered, which could be due to the positioning of the antenna array, indicating that indoor channels are more affected by the array orientation than the position of the system. The joint TX/RX beamforming (eigenvectors) were determined and it was shown that the spatial power characteristics at both carrier frequencies showed a high degree of similarity, suggesting that the multi-path propagation at the two frequencies could be mainly due to specular reflections.

- Maximum entropy approach

  Using the principle of maximum entropy, a new approach in Chapter 6 showed that the eigenvalues of the full covariance matrix is different from the Kronecker model and that there is difference (approximately 5%) in the dominant eigenvalues. Although when comparing the spatial power spectra results (using the ME synthetic full covariance and full covariance matrix of the measured channel), they were close to that using the Kronecker model. The correlation coefficient metric clearly indicated that the the ME model performed better. The study also indicates that the joint RX/TX spatial power spectra at 2.4 GHz and 5.2 GHz are similar, indicating that specular reflections at both the carrier frequencies are similar, hence validating the results obtained for the double directional model in Section 5.3.

7.2 FUTURE RECOMMENDATIONS

Some recommendations to the extension of this work could be:

- Enhanced geometric channel model and joint correlation matrix

  One could extend the work on geometric modelling and incorporate more complex PDFs at the TX and RX. A non-trivial approach would be to reduce the approximations and use numerical techniques to analyze the space-time correlation equations, so that there could be a more accurate representation of the channel. The metrics can then be compared to determine the level of computational complexity versus accuracy. One could also include three dimensional models at the TX and/or RX and extend on the models in the recent work by [104, 105] and compare this with other measured wideband channels.

- Wideband measurement system and analysis

  The present wideband channel sounder could be improved, especially the synchronization system. Various wireless synchronization schemes using for example, Golay Codes or other super orthogonal codes could be implemented. This will be very useful in outdoor measurements as well as where the TX is located at higher altitudes. The comparison of results for different antenna element spacing or geometries would also be useful to further evaluate the impact on the antenna with regards to capacity and spatial power profiles. Further analysis at a lower center frequency like 800 MHz could be useful and of other indoor type environments so as to have a more comprehensive comparative analysis that would provide a useful database for the network planner when making appropriate space-time coding choices.

- Maximum entropy based modelling

  The new modelling approach presented shows potential and has opened up new possibilities for further investigation and analysis. One of the simpler approaches of using linear programming was chosen in order to find the initial guess and then the gradient descent method was used to find optimum values for the system of non-linear equations. Other methods should be investigated to determine the relative accuracies achievable. One could also compare this model to other models such as the Weichselberger [102] and the virtual channel representation [106], both for the capacity predictions and the spatial power representation for this measured environment, as well as other environments, with and without perfect CSI.