Wireless Fading Channels Performance based on Taylor Expansion and Compressed Sensing: A Comparative Approach

Theofanis Xifilidis, Kostas E. Psannis*

Dept. of Applied Informatics

University of Macedonia

Thessaloniki, Greece

Emails:{thxifili@uom.edu.gr, kpsannis@uom.edu.gr}

Abstract- This paper addresses the performance of wireless fading channels, namely, Rayleigh, Rician and Nakagami-m fading channels, taking both noiseless and additive noisy cases into account. The performance metrics investigated are channel capacity, channel model variance estimation, required average description length based on Shannon entropy and relative entropy and Symbol error probability assessment. Taylor expansion is considered as an approximation for the aforementioned metrics together with the Compressed Sensing (CS) compressibility rule. Technical comments are provided for supporting simulation results. Mathematical interpretations are provided to support the simulation results along with feasible applications in 5G wireless systems. Conclusions and future research directions finalize the paper.

Keywords

Compressed Sensing (CS), distributions, fading channels, performance, Taylor expansion

1. INTRODUCTION

The concept of wireless transmission is governed by randomness[1]. The transmitted information through the wireless medium is distorted by additive noise namely Additive White Gaussian Noise (AWGN) a random Gaussian distributed random quantity and fading as a multiplicative effect. The above necessitate the modeling of wireless channels by statistical distributions.

Among fading distributions[2], the Rayleigh fading distribution models the Non-Line-of-Sight channel with multiple scatterers in the medium thus indicative of severe distortion. The Rician distribution, on the other hand, models a channel with a dominant Line-of-Sight component thus a link between transmitter and receiver. As a third fading channel, Nakagami-m fading also accounts for distortion and also approximates the two former distributions by proper parameter selection.

The majority of signal processing applications involve the procedure of estimation, thus, approximating a function, deterministic or random, by a sequence of basis functions with near optimal-minimum error representation. Many such schemes are theoretically formulated by the requirement of infinite number of terms to approach a vanishingly erroneous representation. Polynomial representations are most attractive in various applications.

Taylor polynomial expansions are the best polynomial approximations. This means that if a power series representation is of minimum error, this must be the Taylor expansion of the function to be approximated. Categorized in the class of infinite required terms for zero error, the approximation in practical cases inevitably involve an approximation error, also known as Taylor remainder in the related mathematical literature. The n^{th} order Taylor expansion require the nth times diffferentability of the function to be approached. Being a polynomial, the most important advantage of Taylor polynomial is the ease of mathematical manipulation i.e. elementary integration and differentiation rendering complex integrals calculation with no analytic solution which would otherwise require the application of numerical methods, directly solvable by the respective Taylor polynomial quantity. The convergence of Taylor expansions are point wise or more accurately in a neighborhood around a point called the expansion point as opposed to other infinite order schemes as the Fourier series resulting in interval-wise convergence. Thus, the Taylor polynomial convergence depends on both expansion point and number of terms as powers of the polynomial variable. Different expansion points may require different number of terms to achieve a certain acceptable approximation. Moreover, several terms of the Taylor polynomial may cancel out due to the respective derivative being zero. A pattern resulting in reduced polynomial terms is the even functions class with Taylor representation by even power terms only and the case of odd functions being approximated by Taylor polynomials of odd power terms.

Compressed Sensing (CS)[3] is a rapidly emerging scientific field aiming at significantly reducing computational and implementation complexity caused by traditional sampling and information processing. The mathematical tools that CS uses combine from diverse areas such as statistics, matrix theory, random processes and optimization theory. The fundamental assumption in CS theory and practice is that of sparsity borrowed by matrix algebra indicating the few nonzero elements as opposed to the bulk of zero values. The main concept used in the current paper is compressibility, instead of sparsity. The former sets a predefined threshold relative to elements magnitude thus keeping the values with magnitude above the threshold and discarding the rest.

The concept of Shannon entropy[4] quantifies the uncertainty relative to a random variable with respect to the possible outcomes of the former. Thus entropy is directly applied to random processes of the fading channels. A random variable is always characterized by a nonzero uncertainty while a zero value indicated non-randomness i.e. a deterministic mapping of the variable and its outcomes. A complementary entropy definition also used in Information Theory and Wireless Communications, is the relative entropy quantifying the penalty induced by approximating a supposed true distribution by an approximating one. The sum of the above two entropies results in the average required code description length with respect to a random variable. The above concepts, wireless fading distributions, Taylor expansions and CS together with

Shannon and Relative Entropy, constitute the metrics for assessing performance of the considered fading channels.

2. RELATED WORK AND CONTRIBUTION

Concerning active research of applicability of CS in wireless communications, the authors[5] cover these extensive applications in their review. Relative to 5G communications, channel estimation in Massive MIMO environment[6] is conducted taking channel sparsity into account and solving via CS algorithm. Moreover, Taylor precoding of correlated Massive MIMO is investigated[7] considering low complexity and convergence.

Regarding Taylor approximation of probability density functions, a representative work[8] approximates Normal, Weibull and lognormal densities by a higher degree polynomial based on calculated moments, whereas our work assumes quadratic Taylor polynomial, the first two moments i.e. mean and variance for Rayleigh, Rician and Nakagami-m probability densities. Extending to 5G applications of Taylor expansion, the authors[9] Taylor expand the non-analytic distribution proposing a nonlinear equalization scheme while[10] considers the Taylor expansion throughput in a Device-to-Device interference environment. Furthermore, a nonstationary channel model accounting for Doppler frequency estimation[11] is utilized by Taylor expansion of second degree, as is the case of our paper, to model channel parameters.

Regarding capacity in MIMO Channel via Relative entropy consideration to model mismatch between unknown true distribution and partially known approximating one, the authors[12] solve an optimization problem over distribution and covariance matrix indicating exploitation of correlation while our work involves independent branches and variance for each branch. Noise variance in Massive MIMO system is addressed[13] introducing maximum likelihood estimator verifying the effectiveness for 5G communications. Signal reconstruction[14] defining criteria for distributions to be compressible is investigated, a concept similar to this paper employing compressibility in order to discard small values of channel gains produced by fading distribution. A key element differentiating this work to our paper is the consideration of second moment tending to infinity while the fading cases in our paper always assume finite moments. Additionally, compressibility has also been considered[15] focusing on fourth moment, when the latter is large required multipath components required are in agreement to the number required by CS reconstruction. Finally, in our previous work[16] fading channel coding is investigated employing CS to conduct optimization problem via Lagrange multiplier also solving the inverse distribution identification problem whereas in this paper extended metrics are used to evaluate channel performance via the CS and Taylor approximations.

The contribution of this paper is stated in the following: instead of Taylor expanding the capacity curves or results of estimated variance, which would fail to capture the randomness of the channel gains, or deriving random Taylor coefficients for each fading channel case, we expand via quadratic Taylor polynomial the fading distributions, Rayleigh, Rician or Nakagami-m fading, inferring the same distribution category for the CS-based case. Taylor polynomial is used as the random channel gain generating distribution to assess the wireless fading channel performance just as the exact and CS-based cases, based on capacity, variance, required number of bits for channel coding and symbol error probability estimations. The excellent match of the above calculations to the distribution approximated curves verify the conclusions reached and extend their applicability in the rapidly evolving 5G wireless communications.

3. SYSTEM MODEL AND MATHEMATICAL PRELIMINARIES

3.1 System Model

This paper assumes a single user Single-Input-Multiple-Output system model as shown in Figure 1, hence, a transmitter equipped with a single antenna and a receiver with multiple antennas. The model assumed is simplified not accounting for user mobility i.e. Doppler shift and interference by neighboring transmission. Furthermore, the model assumes that the reception conveys the diversity of independent branches corresponding to the path gain of each receive antenna. Hence, correlation between multipath is not assumed along with inter-element antenna correlation. Moreover, the time coherent model is assumed that is the channel gain over one symbol period is assumed constant.

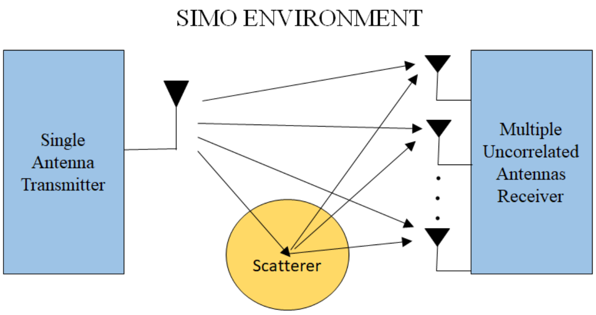


Figure 1 SIMO System Model.

3.2 Mathematical Preliminaries

This section reviews the mathematical tools utilized in this paper.

3.2.1 Fading Distributions

The Rayleigh fading channels is modeled by the following distribution and has the mean and variance below:

$$f_x(x) = \frac{x}{sigma^2} e^{-\frac{x^2}{2sigma^2}} \quad (1)$$

$$mean = sigma\sqrt{\frac{\pi}{2}} \quad (2)$$

$$\sigma^2 = sigma^2\left(\frac{4-\pi}{2}\right) \quad (3)$$

where sigma is the scale parameter by which the first two moments mean and variance are calculated with arbitrary precision since the calculations involve irrational number.

The Rayleigh channels is used in scenario where the medium is filled with multiple scatterers which is result in the multiple paths induced by single or more reflections in the objects. Rayleigh fading assumes severe distortion.

The Rician distribution models wireless channels with a Line-of-Sight component between transmitter and receiver. Its mathematical expression includes apart from the exponent as the Gaussian and Rayleigh distributions, the modified Bessel function of first kind and zero order. The distribution, mean and variance are given below:

$$f_x(x) = \frac{x}{sigma^2} e^{-\frac{(x^2+v^2)}{2sigma^2}} \mathrm{I}_0\left(\frac{xv}{sigma^2}\right) \tag{4}$$

$$mean = sigma\sqrt{\frac{\pi}{2}} * L_{1/2}\left(-\frac{v^2}{2sigma^2}\right) \tag{5}$$

$$\sigma^2 = 2sigma^2 + v^2 - \frac{\pi sigma^2}{2} * L^2_{1/2}\left(-\frac{v^2}{2sigma^2}\right) \tag{6}$$

where sigma and v are parameters to be configured. For zero value of parameter v, the Rician distribution is equivalent to Rayleigh distribution with sigma scale parameter. In the moments expressions $L_{1/2}$ is the Laguerre polynomial with parameter q=0.5.

Finally, the Nakagami-m distribution with the parameter m is also indicative of severe signal distortion where the distribution and first and second moments are given as follows:

$$f_x(x) = \frac{2m^m}{\Gamma(m)\Omega^m} x^{2m-1} e^{-\frac{mx^2}{\Omega}} \tag{7}$$

$$mean = \frac{\Gamma\left(m+\frac{1}{2}\right)}{\Gamma(m)}\left(\frac{\Omega}{m}\right)^{\frac{1}{2}} \tag{8}$$

$$\sigma^2 = \Omega\left(1 - \frac{1}{m}\left(\frac{\Gamma\left(m+\frac{1}{2}\right)}{\Gamma(m)}\right)^2\right) \tag{9}$$

where m controls the shape of distribution and Ω determines the spread. The Nakagami-m distribution can be considered a generalization of the Rayleigh distribution by setting m=1 while an approximation of Rician fading for m>1. The function Γ(.) denotes the gamma function.

Finally, regarding the AWGN distribution the $N(0,1)$ model i.e. zero mean unit variance Gaussian distribution is assumed as given below:

$$f_x(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \tag{10}$$

3.2.2 Taylor polynomial expansions

The approximating Taylor polynomial given a function is given by the power series notation below:

$$T_N(x) = \sum_{n=0}^{N} c_n (x-a)^n = \sum_{n=0}^{N} \frac{f^{(n)}(a)}{n!}(x-a)^n \quad (11)$$

The notation above complies with practical approximations thus necessarily truncating the theoretically requirement of infinite terms for exact approximations. The series coefficients are defined by the value of the n^{th} derivative of the approximated function at the expansion point and then divided by the factorial quantity (n!).

3.2.3 Shannon Entropy and Relative Entropy

The concepts of Shannon and Relative Entropy are fundamental in Information Theory, the former quantifying uncertainty of a random variable and the latter being a means of quantifying the distance or dissimilarity between two distributions. The Shannon entropy and Relative Entropy in a discrete form (a summation assumed instead of a continuous integral) are given below:

$$H(f_x) = -\sum_x f_x \log_2(f_x) \quad (12)$$

$$D(f_x \parallel p_x) = \sum_x f_x \log_2\left(\frac{f_x}{p_x}\right) \quad (13)$$

where f_x is the actual distribution and p_x used in the relative entropy expression is the approximating distribution.

4. SIMULATIONS ANALYSIS

The steps towards the simulations conducted by MATLAB are described as follows.

The initial step is the CS based inferring of the considered fading distribution and the Taylor approximation of the latter.

First, the independent channel gains based on the defined distribution are generated and 50 realizations are carried out producing 50 channel gains each. Then the average channel gains are computed. By employing CS sparsity-enforcing scheme best described as compressibility only the channel gains larger in magnitude than a predefined threshold are kept discarding the rest. These reduced in number channel gains are used to infer the same as originally specified distribution of the same kind. Moreover, the Taylor polynomial approximating the specified distribution is used as the generating distribution producing same amount of channel gains. Based on exact distribution, Taylor polynomial distribution and CS based inferred distribution the capacity of each independent SIMO branch is computed and graphed.

The next step is towards variance estimation based on the first and second order moments. All three cases corresponding to the above distributions are considered providing results by Taylor approximation of the related integrals. It must be noted that irrespective of the feasibility of analytic solutions of the integrals in some cases, the Taylor approximation is applied to all cases for variance estimation. Specifically, the cases of integrals with no analytic solution is where the Taylor expansion offers a tractable and easy to mathematically manipulate approach. The cases accounted for are exact integral calculations, the integral calculation based on the Taylor expansion of the former exact distribution expression and the inferred CS based distribution from the reduced number of channel gains. Proceeding one step further, the required number of bits to express the random channel distribution is calculated being the sum of the Shannon Entropy and the Relative entropy, the first term with respect to the random distribution and the second considering the exact distribution and its Taylor expansion as an approximation or as a second case the CS based inferred distribution as the approximation. The last metric considered is the Symbol Error Probability of the random channel and its comparison to the Taylor expansion and CS inferred distribution.

In all the above investigations, both noiseless case as well as more practical noisy cases are included, the latter expressed by a distribution equal to the convolution of the fading distribution and the AWGN distribution, by assuming independence of fading and noise in the channel.

4.1 Fading Distributions and Channel Capacity

Regarding system model parameters, the SIMO model assumes 50 independent branches thus of transmitted power of 50mW for each branch and a 1GHz bandwidth.

4.1.1 Rayleigh Fading

The first case is Rayleigh fading without noise and Rayleigh fading with additive noise included. The CS based inferred distribution is based on the reduced number of samples with magnitude larger than the average value of the channel gains. Hence, the predefined threshold is set to the average value of channel gains. Regarding the Taylor approximation the quadratic expansion is assumed as the minimum polynomial order capturing curvature. This assumption is based on the convexity of all fading distributions considered. For Rayleigh noiseless, the exact Rayleigh distribution, the CS based inferred distribution and the Taylor approximation are plotted in the following Figure 2:

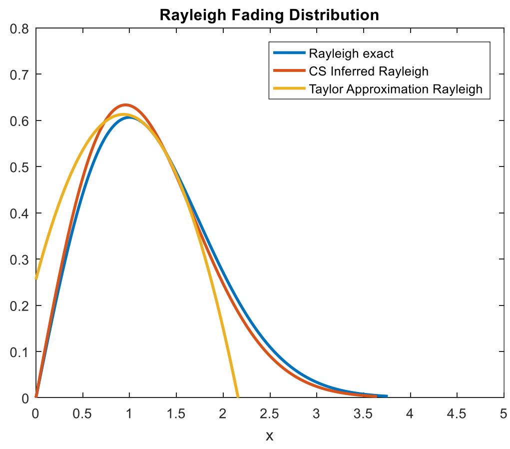


Figure 2 Rayleigh noiseless fading distributions: exact, CS inferred and Taylor approximation.

After running the realizations and computing the average channel gain the capacity with exact Rayleigh fading, Taylor expansion and CS inferred distribution for the noiseless case are depicted in the following Figure 3:

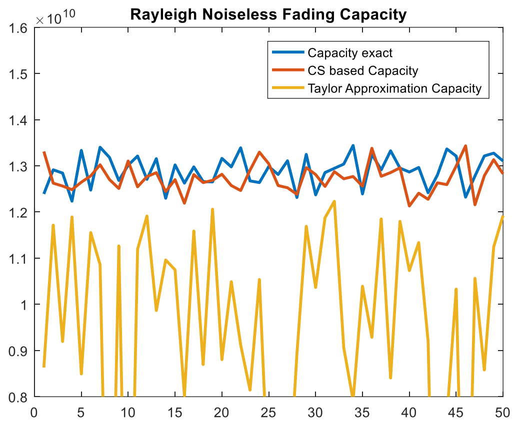


Figure 3 Channel Capacity for noiseless Rayleigh Fading.

Due to the randomness of the channel gains, the capacity preserves the same property. For reaching meaningful results the average values of the three curves is considered for verifying the approximation as well the variation with respect to the mean. The average capacity over the realizations conducted for exact Rayleigh distribution is 12.8Gbps. For CS based inferred distribution the average value was found equal to 12.7Gbps. Finally, for the Taylor approximation the average capacity was 9.82Gbps. The above observations indicate a 0.78% percentage of capacity penalty induced by the compressibility of CS approach equal to 100Mbps and a 23.3% percentage penalty for the Taylor approximated capacity, a quantity of 2.9Gbps. Hence, the two sub-optimal approaches underestimate the capacity. While the CS based inferred distribution case results in near optimal performance, the Taylor approximation approach leads to a significant error in average capacity and also exhibits more variation around its average value. The second case is Rayleigh fading with additive noise included. As in the previous case, the distributions of exact, CS inferred and Taylor approximation Rayleigh fading are shown in Figure 4:

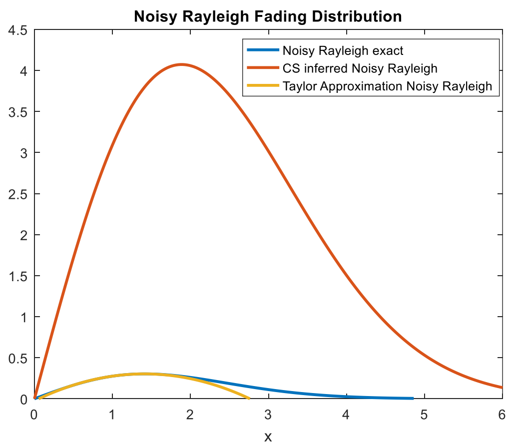


Figure 4 Rayleigh Noisy Fading Distributions: exact, CS inferred and Taylor approximations.

Although in noiseless fading the CS and Taylor approximations provided optimal results the approximated distributions in this case convey the following: the Taylor approximation provides the best fit while the CS inferred distribution results in a distribution with larger variance and hence greater uncertainty for the same interval.

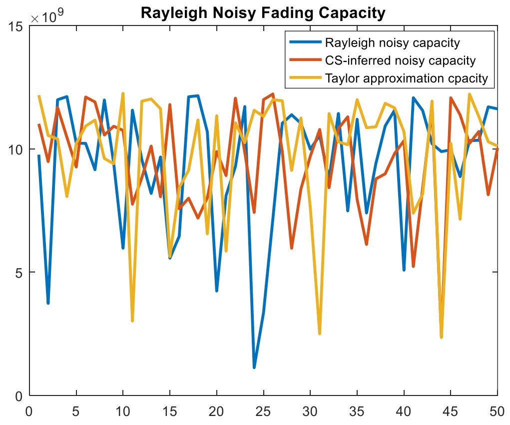


Figure 5 Channel Capacity for Noisy Rayleigh Fading.

The average capacity in Figure 5 for exact distribution equal to the convolution of Rayleigh fading and additive noise was found to be 9.22Gbps. For the CS inferred distribution, the average was found to be 9.44Gbps, 2.3% larger capacity, and for the Taylor approximation it was found equal to 9.48Gbps, 2.7% larger capacity. Hence, the approximation methods resulted to higher average capacity. This is a quite interesting result and can only be justified by the additive noise consideration. As the accurate value was computed by means of the exact Rayleigh noisy distribution, the overestimation observed is that due to noise, a larger interval of samples is needed for the average capacity to settle down and be at most comparable or less than the exact noisy distribution based average capacity. Hence, more samples are needed for the approximation methods to reach a representative value of average capacity. This effect can also be interpreted as increase of variance of the CS distribution compared to the exact average value. Finally, the reduced average capacity compared to the noiseless case is also a direct consequence of the additive noise.

4.1.2 Rician Fading

The second case is the Rician fading distribution both noiseless and noisy cases investigated. The noiseless Rician distribution exact, CS inferred and Taylor approximation are plotted in Figure 6.

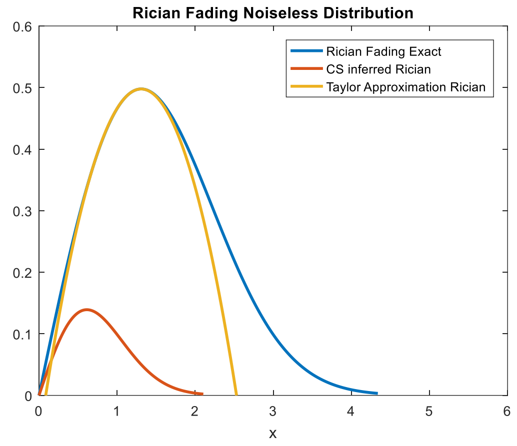


Figure 6 Rician Noiseless Fading Distributions: exact, CS inferred and Taylor approximation.

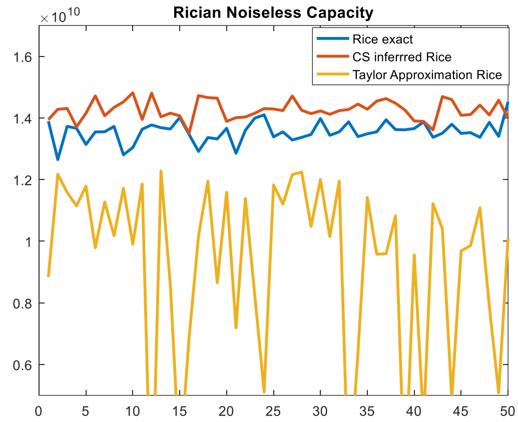


Figure 7 Channel Capacity for Noiseless Rician fading.

The average capacity in Figure 7 by the exact distribution expression was found equal to 13.5Gbps, the CS inferred distribution equal to 14.2Gbps and for the Taylor approximated distribution the average capacity was equal to 10Gbps, inducing a 26% capacity penalty.

The Rician noisy fading distributions are given in Figure 8:

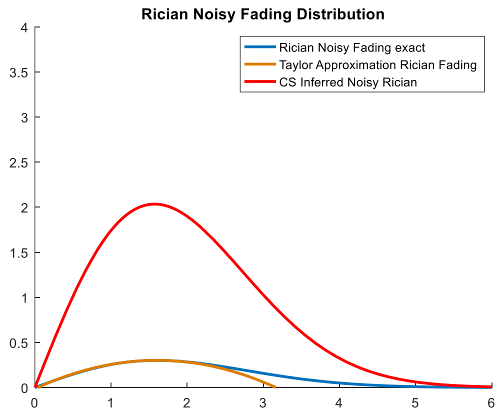


Figure 8 Rician noisy Distributions: exact, CS inferred and Taylor approximation.

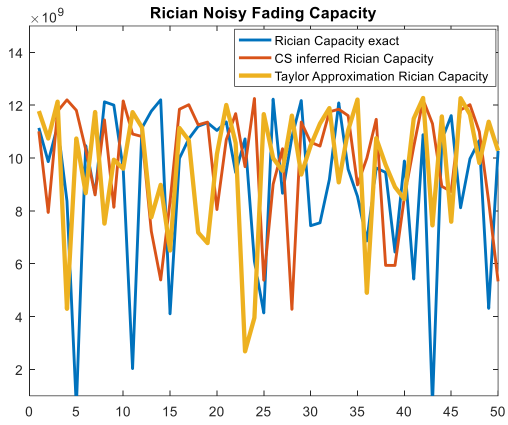


Figure 9 Channel Capacity for noisy Rician fading.

As observed from Figure 9, the average capacity for the exact distribution computation was found equal to 9.78Gbps, for the CS based inferred distribution equal to 8.89Gbps, therefore a 9.1% penalty, and for the Taylor approximated the average capacity was equal to 9.78Gbps, indicating an optimal match related to the Taylor approximation.

The Taylor approximation of the distributions results in close approximation for all cases regardless of the effect of noise and focusing only on the specific distribution. A general remark regarding the CS based result is that there is no significant penalty induced by applying the specific strategy.

4.1.3 Nakagami-m Fading

First, the noiseless Nakagami-m fading case is considered with m shape parameter equal to one and the scale parameter Ω equal to one as well, distributions are given in Figure 10 below:

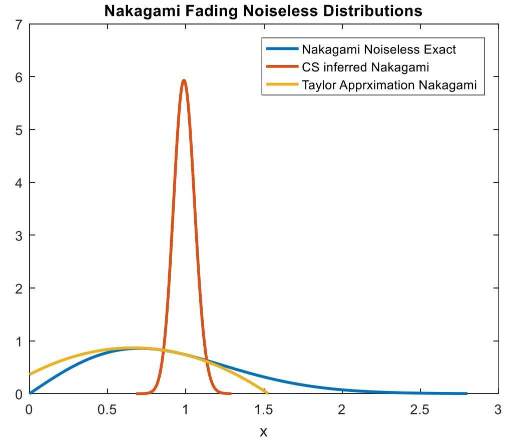


Figure 10 Nakagami-m Fading Noiseless Distributions: exact, CS inferred and Taylor approximation.

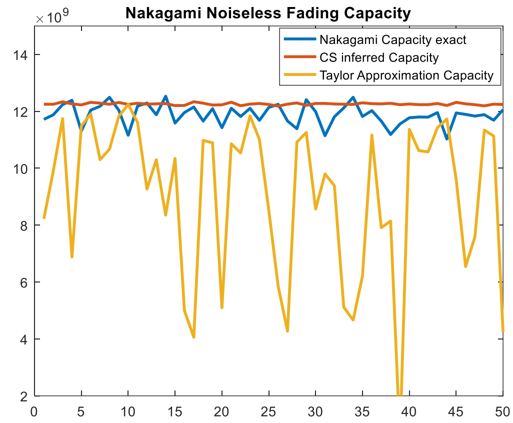


Figure 11 Channel Capacity for Nakagami-m noiseless fading.

The average capacity, Figure 11, based on the exact distribution expression was found equal to 11.9Gbps, while the capacity based on the CS based inferred distribution was found equal to 12Gbps. The capacity based on the Taylor approximated distribution was equal to 9.58Gbps, indicating a penalty of 19.4% in capacity. The CS based result stems from the combined effect of inadequate number of samples to reach a representative mean value along with optimal capacity value. The noisy Nakagami-m fading distributions are given in Figure 12:

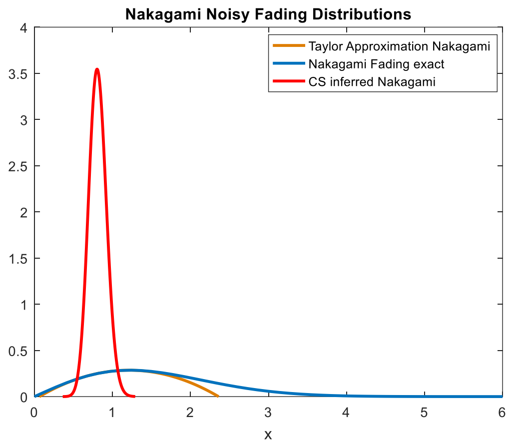


Figure 12 Nakagami-m Noisy Fading Distributions: exact, CS inferred and Taylor approximation.

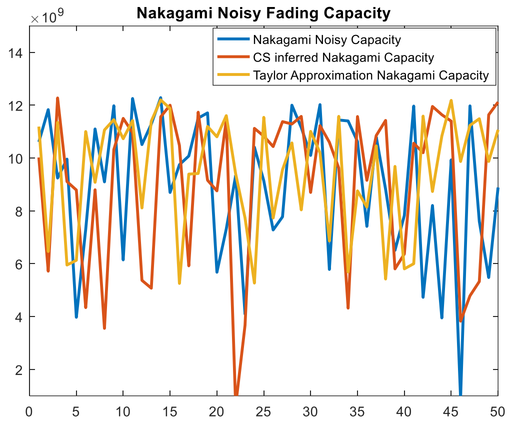


Figure 13 Channel Capacity for Nakagami-m noisy fading.

Based on the average capacity value computations, the average capacity in Figure 13 for the expression was found equal to 8.7Gbps, for the CS inferred distribution the average value was equal to 9.5Gbps and for the Taylor approximated case the average capacity was calculated to be 9.6Gbps.

Finally, the effect of capacity reduction due to additive noise included is evident in all three cases of fading distributions, verifying the accuracy of results in noiseless as well as noisy fading.

4.2 Variance Estimation

This section investigates variance estimation for the three fading channel models considered both for the noiseless as well as the noisy cases. The variance estimation is carried out by means of the equation of computing variance based on the first moment (mean of distribution) and the second moment, the latter two quantities given by the equivalent integrals having the product of independent variable and the corresponding fading distribution as the integrand quantity. The value resulting from these calculations applying the exact distribution expressions, the CS inferred distributions and the Taylor approximated distribution, are then compared to the standard value of the formula for the variance calculation for each distribution. A further note is that the equivalent variance for the noisy cases is equal to the noiseless variance plus one, a direct result from the fading and noise independence assumption and the unit variance noise. For the three fading distributions with noiseless and additive noise included cases, the variances are given in Table I below:

Table 1 Variance Values for fading distributions

Fading Distributions/No noise-Noise included	Noiseless case	Additive noise included
Formula Calculation/Rayleigh fading exact	0.429/0.429	1.429/1.258
Rayleigh CS inferred	0.3935	0.1206
Rayleigh Taylor Approximation	0.2	0.7

Formula Calculation/ Rician Fading exact	0.61/0.61	1.61/1.481
Rician CS inferred	0.1	0.2
Rician Taylor Approximation	0.468	0.9
Formula Calculation/ Nakagami Fading exact	0.214/0.214	1.214/0.946
Nakagami CS inferred	0.005	0.012
Nakagami Taylor Approximation	0.1459	0.476

Based on the results derived by the extensive simulations conducted, the first observation is the perfect match regarding the values from the variance formula for each distribution and the calculation based on the first two moments from the equation involving the related integrals for the noiseless cases. The second observation is the increase in variance for the noise included cases compared to the noiseless cases except for the CS based inferred Rayleigh distribution. Another observation is the mismatch in the noisy case between the independence based calculated variance, equal to the noiseless variance plus one (due to unit variance noise) and the variance calculated by the moment based variance expression. This mismatch is a direct consequence of the concept of independence assumed for the former computation contrary to the latter whereas a correlation assumption is included in the moment integral method. Additionally, the significantly narrow CS based inferred distributions for the Nakagami-m fading cases are verified by the relatively small variance values.

4.2.1 Rayleigh Fading Distribution

In the noiseless case, the variances based on the moment integrals and the independence based are in perfect match. The proximity of the exact and CS-based inferred distribution is also verified by the small deviation of variance values. The slightly lower variance of the Taylor approximation is also verified by Fig.2.

For Rayleigh noisy case, the mismatch of exact moment integral-based capacity compared to the independence based calculated value is evident and previously justified. The variance based on CS-inferred distribution is lower and the Taylor approximation based more closely matches the exact variance as observed from Fig.3. It must be noted that the smaller variance for the CS- inferred distribution is not related to the larger peak value of the distribution but rather on the spread of the curve with respect to its peak value, as observed from the curve.

4.2.2 Rician Fading Distribution

In the noiseless case the moment integral and independence based formula result in the same value. The CS-inferred distribution has an observable lower variance while the closer Taylor approximating curve variance approaches the exact value.

For the noisy case, the mismatch of the two exact variance calculated values is also verified as in the Rayleigh fading case. The CS- inferred distribution is also above the exact distribution curve with significantly lower variance and the Taylor approximation closely matches the exact distribution.

4.2.3 Nakagami-m Fading Distribution

In the noiseless case the moment integral based and independence based calculations indicate a perfect match as in the previous cases. The essential observation for this case is the very narrow distribution curve for the CS-inferred distribution and the observation of the curve being over the exact distribution, which was observed only in the noisy cases for the Rayleigh and Rician fading cases. The significantly small value of the CS-inferred case confirms the narrow shape of the curve, while the Taylor approximation curve produces a more accurate variance with respect to the exact variance value.

In the noisy case, the mismatch between the moment integral based method and the independence based method in terms of variance calculation is also observed. The CS-inferred distribution is very narrow similar to the above Nakagami-m noiseless case indicative of the small variance value, while the Taylor approximation curve provides a closer value to the exact.

4.3 Required Number of bits based on Shannon Entropy and Relative Entropy

In this section the average required number of bits to describe the channel is derived for all fading distribution cases considered above. This number of bits is equal to the Shannon entropy with respect to the exact fading distribution expression plus the relative entropy where the exact distribution expression is used as the reference function and as two separate cases above the approximating distribution is equal to the CS inferred distribution and in the second case the Taylor approximated distribution respectively. The required number of bits for all fading cases are given in Table 2:

Table 2 Required Number of bits for Fading Channels

Fading Distributions/ Entropies	Shannon Entropy exact	Relative Entropy exact vs. CS based inferred/ exact vs. Taylor approximation	Average number of bits required
Rayleigh Noiseless Fading	23.5	3.6/-1.7	28/24
Rayleigh Noisy Fading	17.6	-16.4/ 5.4	18/24
Rician Noiseless Fading	24.4	82.8/17.8	108/43
Rician Noisy Fading	22.6	-29.3/11.4	23/35
Nakagami-m Noiseless Fading	17.8	-70/ -3.2	18/18
Nakagami-m noisy Fading	26	-53/0.86	26/27

Before proceeding to the explanation of the above results regarding the total number of required bits for describing the random fading channels under investigation, some crucial notes related to the justification of the results in this section are necessary.

Although the Shannon entropy is formulated with a sum including the logarithm of values smaller than unity, the minus sign is necessary in order for the Shannon entropy to be a positive quantity. This entropy quantifies the number of bits required due to the uncertainty of the respective random variable.

On the other hand, the relative entropy comprises of a sum including the logarithm of the ratio of the exact distribution values to the approximating distribution values, thus, quantifying the distribution mismatch occurring due to this approximation. This mismatch is thus translated to the penalty of additional bits required to describe the channel.

Regarding the results of the relative entropy values for the fading distributions considered, some cases derived negative values of the relative entropy. Initially, this is completely in contrary to the definition of the relative entropy. However, the results produced can be justified. The relative entropy involves approximation of a distribution by another such that the quantity inside the logarithm is greater than unity. Hence, the relative entropy is based on the property of approximating a given distribution by one that has smaller uncertainty, hence, the additional bits required to describe the exact distribution.

Given the fading distributions in the cases of negative relative entropy, the approximating fading distribution is above the exact distribution hence the ratio inside the logarithm is smaller than unity, and the negative valued relative entropy is justified. In terms of interpreting the result of the required bits to describe the channel, these negative values are regarded as zero additional bits required. Hence, the channel is fully described by the derived positive Shannon entropy. The issue of the actual effect of negative relative entropy as a means of reducing the overall required bits for channel coding is left as future research.

A remark concerning the Taylor approximation curve is the following: although the exact fading is always positive valued the Taylor approximation curve, quadratic as already stated, approaches the value at most at the expansion point and then deviates as the neighborhood region expands. Thus, after crossing the horizontal axis, it takes negative values. Hence in order to regard this polynomial curve as a distribution we impose a constraint that the curve considered is in the positive valued interval.

4.3.1 Rayleigh Fading Distribution

In the noiseless case, the required number of bits are 24 due to Shannon entropy and, for the CS-inferred case, 4 additional bits a value verified by the close match of the two distribution curves. The negative valued relative entropy for the Taylor approximation case is the combined effect of the proximity of the two curves and the Taylor approximation curve being slightly above the exact curve thus being the factor that contributes to the negative overall value. As a result, the required bits for the CS-inferred case is equal to 28 bits. For the Taylor approximation curve, 24 bits (Shannon entropy) are required.

In the noisy case, the Shannon entropy contributes to 18 bits. For the CS-inferred case, the distribution curve is above the exact curve producing negative relative entropy as justified above, whereas for the Taylor approximation case, 6 additional bits are required. Thus, for the CS-inferred case 18 bits are required and for the Taylor approximation 24 bits are required.

4.3.2 Rician Fading Distribution

In the noiseless case, the Shannon entropy dictates the need for 25 bits. For the CS-inferred case, 83 bits are needed indicative of the significant mismatch between the two curves and for the Taylor approximation 18 bits additionally are required. Hence for the CS case overall 108 bits are required and for the Taylor approximation case a total of 43 bits are required for channel description.

In the noisy case, 23 bits are required due to Shannon entropy. For the CS inferred case, the distribution curve being above the exact distribution results in negative relative entropy. For the Taylor approximation, 12 additional bits are required. Hence, for the CS case 23 bits are sufficient and for the latter case, 35 bits are required.

4.3.3 Nakagami-m Fading Distribution

In the noiseless case, 18 bits are required by the Shannon entropy. For the CS-inferred case and Taylor approximation cases, both derive negative values for relative entropy. The former is due to the narrow distribution above the exact curve and the latter due to the part of the respective Taylor approximation curve slightly above the exact curve. Hence, for both cases, 18 bits are required.

Finally, in the noisy Nakagami-m case, the Shannon entropy derives an amount of 26 bits required, whereas for the CS-inferred case, the negative valued relative entropy is justified similarly to the Nakagami-m noiseless case. For the Taylor approximation, 1 additional bit is required. Hence, for the CS-inferred case and Taylor approximation case 26 bit and 27 bits are required, respectively.

4.4 Wireless fading performance analysis for Fading Distributions

The performance of the fading channels considered with the exact, CS-inferred and Taylor approximation distributions assumed is quantified below by the Symbol Error Probability for each fading case.

4.4.1 Rayleigh Fading noiseless-noisy cases[17]

Firstly, for the noiseless case, the error curves of exact and CS-inferred cases, are similar. This constitutes the first essential conclusion being that with the CS compressibility principle, taking only the largest in magnitude channel gains into account, does not sacrifice performance. For the Taylor approximation case higher error probability is observed, compatible to the exact distribution approximation by the Taylor polynomial curve.

For the noisy case, the CS-inferred curve is observed slightly with a smaller error probability. This is a combined effect of no performance degradation and the issue of how representative the error curve is for a certain number of samples, or, in other words how many samples are required to reach an accurate performance result. The Taylor approximation error curve is also indicative of a significantly smaller penalty compared to the noiseless case. This can be directly related to the slightly closer mean capacity value for the noisy case, compared to the noiseless capacity case. This smaller penalty can be directly linked to the closer match of the average capacity for the Rayleigh fading with additive noise included compared to the noiseless Rayleigh fading. Symbol error probability curves for Rayleigh fading are plotted in Figure 14:

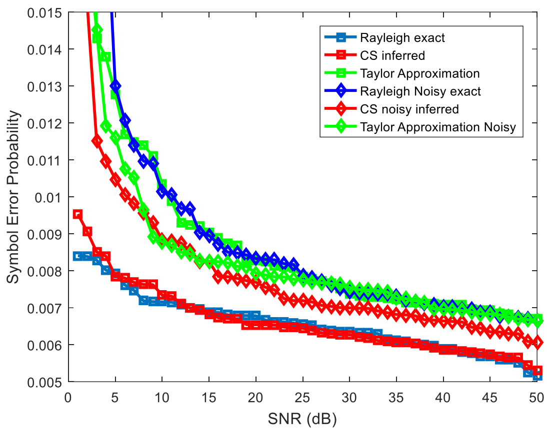


Figure 14 Symbol Error Probability for Rayleigh fading noiseless –noisy cases.

4.4.2 Rician Fading noiseless-noisy cases[18]

In the noiseless case, the result of no performance degradation for the exact and CS-inferred case is observed as in the Rayleigh distribution. The Taylor approximation case results in this case as a performance penalty.

In the noisy case, the CS-inferred error curve is observably below the exact curve and can be justified by the need for more samples to approach the exact curve as well as no performance loss. The Taylor approximation error curve for the noisy case is almost identical with the exact distribution error curve verifying the close approximation in this case. Symbol error probability curves for Rician fading case are given in Figure 15:

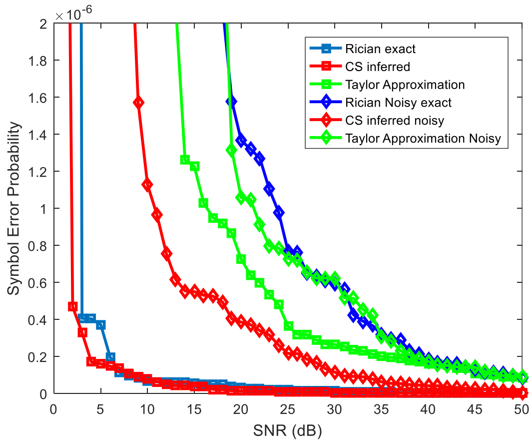


Figure 15 Symbol Error Probability for Rician fading noiseless noisy cases.

4.4.3 Nakagami-m Fading noiseless-noisy cases[19]

In the noiseless cases, the CS-inferred error curve depicts a performance degradation which is anticipated by considering the respective distribution curve derived in the previous section being above the exact. The Taylor approximation curve indicates a small performance penalty compared to the exact distribution error curve. Symbol error probability curves are given in Figure 16:

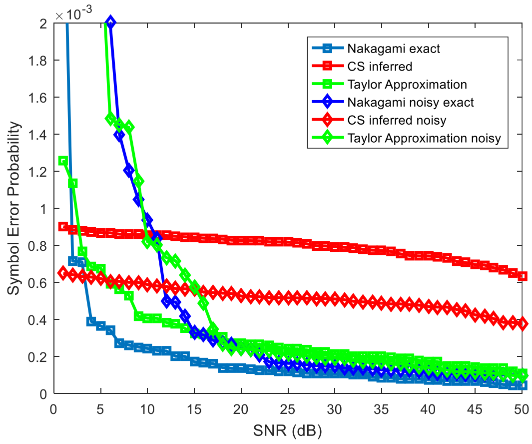


Figure16 Symbol Error Probability for Nakagami fading noiseless noisy cases.

The same observation holds for the CS-inferred error curve in the noisy case, i.e. a performance degradation compared to the exact. Finally, the Taylor approximation error curve is almost identical to the error curve of the exact distribution.

As general remarks for system performance based on symbol error probability, the Nakagami-m is shown to relate to an error probability lower than Rayleigh fading and higher compared to Rician fading. Thus, it is proven to be a compromise between severe scattering assumption and Non-Line-of-Sight conditions for Rayleigh fading and the assumption of a dominant Line-of-Sight existence between transmitter and receiver in the Rician fading cases. Mathematically, this is justified by the fact that Nakagami-m distribution can efficiently model the Rayleigh and Rician fading distributions by proper parameter selection. The performance degradation due to additive noise included is verified for all fading cases.

5. MATHEMATICAL INTERPRETATIONS

Regarding the convexity assumption which is the prerequisite for applying the approximations considered several remarks are important.

The Rayleigh distribution is formulated by an exponential decaying factor, a Gaussian bell curve and a linear term by multiplication. The result is the known non-symmetric convex curve of the Rayleigh distribution. For the Rician distribution the formula is based on the components of the Rayleigh distribution multiplied by a modified zero order Bessel function term, the latter being monotonically increasing similar, in concept, to the linear term. Thus the result is concluded that the exponential decaying factor and the latter terms determine the non-symmetric shape of the Rician fading distribution. The above observations also hold for the Nakagami-m case.

The above verify that the distribution curves are convex, hence the peak value is the unique extremum as for increasing values the distributions decay to zero. This property is the assumption of convexity that verifies the feasibility of the CS-based compressibility approximation method. Moreover, the Taylor approximation polynomial is assumed to be second degree thus the minimum degree for convexity and capturing of the curvature of the fading distributions. For the CS method terms, the above cancels any possibility of additional local extrema justifying the latter method for approximation.

Generalizing the use of Taylor approximation for fading channels, the fading distribution shape for all fading cases are contemplated along with the nth degree polynomial for approximation. Thinking of distributions as the underlying mathematical function, the remark of the zero of this function at the value of zero of the horizontal axis is noted along with a smooth decay of the right side to zero as the x-axis increases in value. Hence, the main advantage of Taylor expansion being the straightforward analytic calculation of the variance integrals, the entropy calculations and probability calculations with the respective integrals is traded for a truncated region of the approximated distribution defined by the extent of the neighborhood in which the Taylor polynomial converges. This is optimized by increasing the degree of the Taylor polynomial.

Hence, for practical cases, the truncation of the distribution neglecting the decaying distribution tail can be achieved by the corresponding Taylor polynomial, the only tradeoff being the increase in complexity as the Taylor polynomial degree increases.

In terms of most accurate approximation, Taylor polynomial is more accurate being a property of increased convergence inside a defined interval while CS inferred case results in a distribution of the same kind but with different parameters. Further defining the above mentioned complexity tradeoff, CS inferred related case trades no performance degradation with closer approximation by Taylor expansion.

6. APPLICATIONS TO 5G COMMUNICATIONS

The graphs of the fading distributions i.e. exact distribution, CS inferred distribution and Taylor approximated distribution along with the rest of the results including variance estimation, entropy based required number of bits and Symbol Error Probability are sufficient to completely describe the fading channel.

The capacity results are enough to provide an upper bound on the achievable rate through the channel as fundamentally stated by Shannon. The Symbol Error probability results account for estimating the fraction of transmitted bits anticipated to be erroneously received by the wireless receiver. Proceeding further, the variance of each distribution indicating the spread of the values under the curve with respect to the peak, is also a measure of uncertainty. Along with the latter, the required number of bits indicates the uncertainty of the channel distribution as well, offering insight into the approximation properties of the distributions that enter the equation as approximating ones.

A crucial remark extending the beneficial approximation of CS based distribution is that besides no performance degradation, hence, no error probability increase is that due to the negative relative entropy no additional bits are required apart from those dictated by the Shannon entropy. Hence, no performance degradation comes with no complexity increase in the more general context of variable length channel coding.

There are numerous benefits from applying the CS and Taylor approximation in wireless communication design aspects particularly in the up-to-date fifth generation (5G) communications systems.

The most important benefit of Taylor expansion is the transformation of complex expressions and integrals from a non-closed form where analytic solution is not available and numerical methods are the only methods applicable, to a polynomial expression enabling differentiability and integration, the former being generally more straightforward compared to the latter.

Examples of the feasibility of Taylor expansion are the derivations of error probability for various channels and modulation schemes, interference management in 5G cellular networks, channel estimation/equalization methods as well as resources allocation in terms of time, frequency, code or power.

In clustered MIMO channels where clusters of multipaths are characterized by a common delay, channel estimation expressions may be modeled by a Taylor expansion of arbitrary degree with expansion point equal to the cluster delay value. Another case of Taylor expansion applicability is communication hardware design, namely power amplifier characteristic the latter being nonlinear in the case of excessive reception power. Both the second or higher Taylor polynomial degree can be used to model nonlinearity. The linear term can also be of interest in case the power amplifier moves away from saturation. Consequently, the Taylor expansion point can also follow certain threshold value crossings in specific cases or even zero crossings if a sufficiently high Taylor polynomial order is used that has the property of its roots as a polynomial being the zeros of the nonlinear curve. Moreover, in waveform signaling, in cases other than the Dirac impulse function in the case of static channels that involve the channel gain for each transmission interval, Taylor expansion may be used to expand the finite duration waveform in the expression for the received symbol.

Regarding CS, though significantly extended in wireless communications literature in the areas of channel estimation, spectrum sharing in cognitive radio as well as distributed system performance, there are numerous areas where CS may prove to be beneficial due to the aforementioned computational and implementation complexity alleviation. The verified conclusion reached in this paper is the decreased complexity along with no performance degradation.

In overall, this paper providing insight by approximating fading distributions and with the aforementioned application to 5G communication does not proceed to channel estimation and assesses performance based on CS and Taylor expansion. Thus, it is closer to the concept of estimation rather than signal reconstruction requiring additional analysis with increased complexity.

7. CONCLUSIONS AND FUTURE WORK

In this paper, wireless fading channel performance is evaluated based on CS approximation and Taylor expansion for the Rayleigh, Rician and Nakagami-m distributions, respectively. The three respective distribution curves i.e. exact, CS inferred and Taylor approximation are derived for each fading channel considering both noiseless and additive noise included cases. The variances for each fading channel are derived based on the aforementioned approximations. Required number of bits for the fading channels based on Shannon entropy and relative entropy calculations are derived while symbol error probabilities are derived evaluating fading channel performance. Technical and mathematical interpretations are provided and potential applicability cases in 5G communication systems are pointed out.

As future research, the fully formulated CS optimization problem for the communication system performance while applying optimal algorithms from CS literature in order to optimize performance with negligible complexity increase is an interesting extension. Apart from the convexity prerequisite for applying CS linear programming for each case as is the case of this paper, an interesting extension encompassing are nonconvex problems with CS enabling their treatment as convex with acceptable optimization results.

Taylor approximation may also be used by relying on sufficiently high order of differentiabilty given specific formulas that correspond to performance metrics. The convenient properties of integration and differentiability as well as the Taylor approximation of complex functions in terms of simpler ones from which they may be derived render this issue a significant challenge for polynomial representation of wireless communication metrics.

DATA AVAILABILITY STATEMENT

The authors confirm that there is no data sharing for this manuscript.

FUNDING

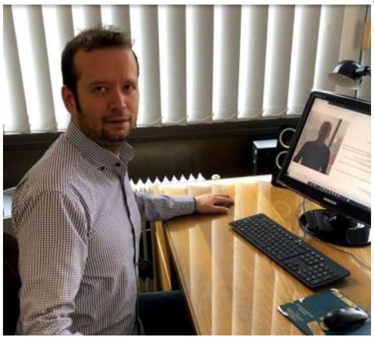
No Funding.

REFERENCES

1. Goldsmith A. Wireless Communications, Cambridge: Cambridge University Press; 2005. 571p.
2. Pätzold M. Mobile Fading Channels. New York: John Wiley & Sons; Ltd. 2002. 430p.
3. Donoho D.L. Compressed Sensing. IEEE Trans. on Information Theory 2006;52:1289-1306.
4. Cover T.M, Thomas J.A. Elements of Information Theory, 2nd . New York: John Wiley & Sons; Ltd. 2006. 774p.
5. Choi JW, Shim B, Ding Y. Compressed Sensing for Wireless Communications: Useful Tips and Tricks. IEEE Comm. Surveys and Tutorials thirdquarter 2017;19:1527-1550.
6. Pramanik A, Maity SP, Farheen Z. Compressed Sensing Channel Estimation in Massive MIMO. IET Commun. J 2019;13:3145:3152.
7. Zhang W, De Lamare RC, Pan C, et al. Correlation-driven Optimized Taylor Expansion Precoding for Massive MIMO Systems with Correlated Channels. In: International Conference on Communications. Paris: IEEE; 2017:1-6.

8. Munkhammar J, Mattsson L, Rydén J Polynomial Probability Distribution Estimation Using the Method of Moments. Plos One J 2017;12:1-14.
9. Li B, Zhao C, Sun M et al. A Bayesian Approach for Nonlinear Equalization and Signal Detection in Millimeter-wave Communications. IEEE Trans. on Wireless Communications 2015;14:3794-3809.
10. Lv S, Xing C, Zhang Z, Long K. Guard Zone Based Interference Management for D2D-Aided Underlaying Cellular Networks. IEEE Trans. on Vehicular Technology 2017;66:5466-5471.
11. Li W, Chen X, Zhu Q et al. A Novel Segment-based Model for Non-Stationary Vehicle-to-Vehicle Channels with Velocity Variations. IEEE Access 2019;7:133442-133451.
12. Charalambous C.D, Denic S.Z, Constantinou C. Information Capacity of MIMO Channels with Relative Entropy Constraint. In: International Symposium on Information Theory Seattle:IEEE; 2006:876-880.
13. Iscar J, Güvenç İ, Dikmese S, Rupasinghe N. Efficient Noise Variance Estimation Under Pilot Contamination for Massive MIMO Systems. IEEE Trans. on Vehicular Technology 2018;67:2982-2996.
14. Gribonval R, Cevher V, Davies M.E. Compressible Distributions for High-Dimensional Statistics. IEEE Trans. on Information Theory 2012;58:5016-5034.
15. Gómez-Cuba F, Goldsmith A.J. Compressed Sensing Channel Estimation for OFDM with Non-Gaussian Multipath Gains. IEEE Trans. on Wireless Communications 2020;19:47-61.
16. Xifilidis T, Psannis K.E, Fading Channel Coding based on Entropy and Compressive Sensing. In: 3rd World Symposium on Communication Engineering Thessaloniki:IEEE; 2020:44-48.
17. Proakis JG. Digital Communications, 5th Edition. New York: Mc-Graw Hill International Edition; 2008. 1170p.
18. Chandra A, Bose C. Error Probability of Coherent Modulations in Rician Fading Channel. International J of Interdisciplinary Telecommunications and Networking 2009;1:16-27.
19. Salahat E, Abualhaol I. Generalized Average BER Expression for SC and MRC Receiver over Nakagami-m Fading Channels. In: 24th Annual International Symposium on Personal, Indoor and Mobile Radio Communications London:IEEE; 2013:3360-3365.

Theofanis Xifilidis received the Diploma of Electrical and Computer Engineering from the Aristotle University of Thessaloniki, Greece. He received his Master's Degree of Electronics and Radioelectrology from the Physics Department Aristotle University of Thessaloniki, Greece. He is currently pursuing his P.H.D. in the Department of Applied Informatics, University of Macedonia, Thessaloniki, Greece. His research interests include Applied Mathematics, Compressive Sensing, Signal Processing, Information theory, wireless communication systems and MIMO wireless technology.



Konstantinos E. Psannis was born in Thessaloniki, Greece. He is currently an Associate Professor in Communications Systems and Networking at the Department of Applied Informatics, School of Information Sciences, University of Macedonia, Greece, Director of Mobility2net Research & Development & Consulting JP-EU Lab and member of the EU-JAPAN Centre for Industrial Cooperation. Konstantinos received a degree in Physics, Faculty of Sciences, from Aristotle University of Thessaloniki, Greece, and the Ph.D. degree from the School of Engineering and Design, Department of Electronic and Computer Engineering of Brunel University, London, UK. From 2001 to 2002 he was awarded the British Chevening scholarship. The Chevening Scholarships are the UK government's global scholarship programme, funded by the Foreign and Commonwealth Office (FCO) and partner organisations. The programme makes awards to outstanding scholars with leadership potential from around the world to study at universities in the UK. Dr. Psannis' research spans a wide range of Digital Media Communications, media coding/synchronization and transport over a variety of networks, both from the theoretical as well as the practical points of view. His recent work has been directed toward the demanding digital signals and systems problems arising from the various areas of ubiquitous big data/media and communications. This work is supported by research grants and contracts from various government organizations. Dr. Psannis has participated in joint research works funded by Grant-in-Aid for Scientific Research, Japan Society for the Promotion of Science (JSPS), KAKENHI Grant, The Telecommunications Advancement Foundation, International Information Science Foundation, as a Principal Investigator and Visiting Consultant Professor in Nagoya Institute of Technology, Japan. Konstantinos E. Psannis was invited to speak on the EU-Japan Co-ordinated Call Preparatory meeting, Green & Content Centric Networking (CCN), organized by European Commission (EC) and National Institute of Information and Communications Technology (NICT)/ Ministry of Internal Affairs and Communications (MIC), Japan (in the context of the upcoming ICT Work Programme 2013) and International Telecommunication Union. (ITU-founded in 1865), SG13 meeting on DAN/CCN, Berlin, July 2012, amongst other invited speakers. Konstantinos received a joint-research Award from the Institute of Electronics, Information and Communication Engineers, Japan, Technical Committee on Communication Quality, July 2009 and joint-research Encouraging Prize from the IEICE Technical Committee on Communication Systems (CS), July 2011. Dr. Psannis has more than 60 publications in international scientific journals and more than 70 publications in international conferences. His published works has more than 3000 citations (h-index 25, i10-index 51). Dr. Konstantinos E. Psannis has been included in the list of Top 2% Scientists in the world (prepared by Stanford University USA, October 2020) (https://lnkd.in/dhSwdgB). Dr. Psannis supervises a post-doc student and seven PhD students. Prof. Konstantinos E. Psannis serving as an Associate Editor for IEEE Access and IEEE Communications Letters. He is Lead Associate Editor for the Special Issue on Roadmap to 5G: rising to the challenge, IEEE Access, 2019. He is a Guest Editor for the Special Issue on Compressive Sensing-Based IoT Applications, Sensors, 2020. He is a Guest Editor for the Special Issue on Advances in Baseband Signal Processing, Circuit Designs, and Communications, Information, 2020. He is a

Lead Guest Editor for the Special Issue on Artificial Intelligence for Cloud Based Big Data Analytics, Big Data Research, 2020. He is TPC Co-Chair at the International Conference on Computer Communications and the Internet (ICCCI 2020), Nagoya Institute of Technology Japan, ICCCI to be held in 2020 June 26-29 at Nagoya, Japan, and Conference Chair at the World Symposium on Communications Engineering (WSCE 2020- http://wsce.org/) to be held at University of Macedonia, Thessaloniki, Greece, October 9-11, 2020.