IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 10, 2019, accepted April 22, 2019, date of publication May 7, 2019, date of current version May 21, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2915099

Power Allocation in 5G Wireless Communication

ZHANGLIANG CHEN, (Student Member, IEEE), AND QILIAN LIANG, (Fellow, IEEE)
Department of Electrical Engineering, The University of Texas at Arlington, Arlington, TX 76010, USA

Corresponding author: Qilian Liang (liang@uta.edu)

This work was supported in part by the NSFC under Grant 61771342, Grant 61731006, Grant 61372097, and Grant 61711530132.

ABSTRACT Granger Causality analysis originated in the field of econometrics is used as a time series analysis tool based on vector auto-regression, and its phased generalized transfer entropy (TE), which is based on conditional co-information in information theory, has been widely used in data analysis in recent years. In this paper, we forecast the Fifth-Generation (5G) channel based on the Granger causality and transfer entropy, then use the water filling algorithm to allocate power for the forecasted 5G channel. In the first part of the paper, we use the Granger causality test to verify the Granger causality correlation of two random 5G channels and ensure that the two channels can be forecasted using the Transfer Entropy method. In the second part, we use transfer entropy to forecast two channels and verify the accuracy of the forecasted channels using Root Mean Square Error (RMSE) and Cramer-Rao Lower Bound (CRLB). Finally, we use the Inverse Water-Filling (IWF) algorithm to perform the power allocation for the forecasted channels and compare it with the Equal Gain (EG) algorithm. The simulations further validate our theoretical results.

INDEX TERMS 5G channel forecasting, transfer entropy, Granger causality, power allocation, inverse water filling.

I. INTRODUCTION

The 5G is a new generation of mobile communication system developed for the demand for mobile communication after 2020 [1]. According to the development law of mobile communication, 5G will have ultra high spectrum utilization and energy efficiency, and increase the magnitude or higher of the 4G mobile communication in terms of transmission rate and resource utilization. [2]. 5G will be an important part of the network society and will help realize the vision of unlimited access to information and data for anyone, anything, anytime, anywhere [3], [4]. Power efficiency has long played an important role in mobile communication devices. The high power efficiency of the device extends battery life and has always been an important element of the revolution in mobile communications [5]. However, the need for high energy efficiency has also become a key point in network infrastructure. Although the amount of communication traffic and users has increased significantly, reducing the overall power consumption of the network is still our goal. The main reasons are as follows:

1) High network power efficiency is the key to reducing operating costs, and is the driving force for better node and network deployment, thereby reducing the total investment cost [6].

2) High network power efficiency is part of the overall goal of the operator and aims to provide wireless access in a way that is sustainable and resource efficient [7], [8].

The 5G power is unevenly distributed. Even in the same area, there is a large spatial difference in 5G power distribution [9]. Therefore, even in dense urban areas with high traffic loading, there is a specific place for low traffic loading [10]. At present, more and more attention is paid to energy efficiency and energy consumption. The implementation of network function should not consume energy excessively. 5G should be able to support lower power consumption and become a greener and environmentally friendly mobile communication network [11], [4]. Therefore, the high energy efficiency goal of reducing power consumption is a key requirement of 5G. The prominent contribution of the paper is applying the TE method on the simulated actual 5G channel to forecast the channel coefficient, then in order to achieve high power efficiency, the paper utilized IWF algorithm to optimize the energy distribution of the forecasted 5G channel based on transfer entropy, compared to the traditional EG distribution method.

The rest of this paper is organized as follows. In Section II, we introduce the concept of Granger Causality and test for 5G Channels. The forecasting is carried out on the basis of the fact that the two channels have correlation and in Section III, we propose the 5G channel forecasting based on transfer

The associate editor coordinating the review of this manuscript and approving it for publication was Tawfik Al-Hadhrami.

entropy. In Section IV, we propose the RMSE and CRLB of forecasted channels, and the comparison between the RMSE and the CRLB for the forecasted channel and the real channel ensures a lower forecasting error. In Section V, we propose the power allocation based on water-filling algorithm. In the process of power allocation, we compare the efficiency of power allocation using the traditional method of equal gain algorithm and the water-filling algorithm. In section Vi, we perform simulations and Performance Analyses. Finally, we conclude this paper and discuss further research directions in Section VII.

II. INTRODUCTION TO GRANGER CAUSALITY AND TEST FOR 5G CHANNELS

In economic analysis, it is often necessary to judge the causal relationship between economic variables. Using statistical inference methods to derive empirical judgments of causal relationships between variables from actual observation data, may be an effective method for causality testing. The Granger Causality test can statistically test the causality between variables. For things that have unclear causal relationships, this method can be used to perform statistical tests [12].

In 1969, Granger proposed a definition of causality from the perspective of econometrics: With two time series x_t, y_t, consider the linear projection of x_t on the past values of x and y:

$$x_t = \sum_{j=1}^{\infty} h_j x_{t-j} + \sum_{j=1}^{\infty} v_j y_{t-j} + \epsilon_t. \tag{1}$$

where, for any positive integer k, $E\epsilon_t x_{t-k} = E\epsilon_t y_{t-k} = 0$ [13], ϵ_t is noise.

If for a given past value of all x, the past value of y contributes to predicting x, ie. there is at least one j_0 such that $vj_0 \neq 0$, then the variable y is the Granger sense reason for x.

According to this definition, in 1972 Sims proposed a proposition that there is no causal relationship which is, Let (x_t, y_t) be a zero-mean joint covariance stationary sequence, then y is not a necessary and sufficient condition for the Granger's cause of x [5]. There exists a vector moving average representation of the lower triangle:

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} a(L) & 0 \\ b(L) & h(L) \end{pmatrix} \begin{pmatrix} \epsilon_t \\ u_t \end{pmatrix} \tag{2}$$

where ϵ_t and u_t are zero-ranging sequence-independent process, and for any integer $t, s, E\epsilon_t u_s = 0$, $a(L), b(L), h(L)$ is the non-negative exponent of L Polynomials on the side, ie., $a(L) = \sum_{j=0}^{\infty} a_j L^j, b(L) = \sum_{j=0}^{\infty} b_j L^j, h(L) = \sum_{j=0}^{\infty} h_j L^j$, in which L is a delay operator defined by $Lx_t = x_t - 1$.

If y is not the Granger cause of x, (2) holds. Rewrite (2):

$$\begin{aligned} x_t &= a(L)\epsilon_b \\ y_t &= b(L)\epsilon_b + h(L)u_t \end{aligned}$$

From the joint covariance stationarity of (x_t, y_t), the inverse polynomial $a^{-1}(L)$ of $a(L)$ exists, and it is non-negative power-universal to the delay operator L.

Let $b(L) = b(L)a^{-1}(L)$, $e_t = h(L)u_t$, and then there are:

$$\begin{aligned} y_t &= b(L)a^{-1}(L)x_t + h(L)u_t \\ &= d(L)x_t + e_t \end{aligned} \tag{3}$$

This formula shows that the regression residual of y_t which is e_t on the current and past x (ie $x_t, x_{t-1}, x_{t-2},$) is not related to the future x_t [14]. In other words, given the current and past x_t, the future x_t does not affect y_t, i.e., (3) is a representation under the condition that y has no feedback effect on x. Observe that Granger's Causality definition assumes that future events cannot cause current or past events. Therefore, the real meaning of Granger's Causality is the 'preceding' relationship in time, not the causality in the usual sense.

According to his above proposition, Sims proved a theorem that facilitates Granger Causality test as follows [15]. Assume that (x_t, y_t) is a zero-mean joint covariance stationary sequence, ϵ_t is a white noise sequence, considering the linear projection of y_t over the whole x process:

$$y_t = \sum_{j=-\infty}^{\infty} b_j x_{t-j} + \epsilon_t$$

where, for any integer j, $E\epsilon_t x_{t-j} = 0$. Then y is not Granger's cause of x, i.e., in (1), for any j, $v_j = 0$, if and only if it is an arbitrary negative integer j, $b_j = 0$.

Granger Causality can be tested using the following metrological methods. Let x_t, y_t be the covariance stationary sequences, set up a regression model of x_t for lags of y and x:

$$x_t = c + \sum_{i=1}^{n} h_t y_{t-i} + \sum_{j=1}^{n} a_j x_{t-j} + \epsilon_t \tag{4}$$

where, c is constant.

Among them, the choice of lag period n is relatively arbitrary.Then the judgment that 'y is not the cause of x' is equivalent to performing an F-test on the null statistical hypothesis:

$$H_0 : h_1 = h_2 = \ldots h_n = 0$$

Let SSR_1(Residual Sum of Squares) and SSR_0 represent the residual sum of squares of regression model(4) and the model when the null hypothesis H_0 holds. Then, the test statistic

$$F = \frac{(SSR_0 - SSR_1)/n}{SSR_1/(N - 2n - 1)}$$

follows the F-distribution with the first degree of freedom n and the second degree of freedom $N - 2n - 1$ under the condition that H_0 holds. Where N is the number of sample data.

When the value of the above F statistic is greater than the critical $F_\alpha(n, N - 2n - 1)$ of the F distribution below the significance level $1 - \alpha$, y can be considered as the Granger cause of x under the confidence of $1 - \alpha$.

We perform a Granger causality test on two random and relatively independent 5G channels in order to verify the

existence of causality between the two channels. If there are two channels without causality, in other words, the change of channel 1 has no effect on the change of channel 2 or the change of channel 2 has no effect on the change of channel 1, then we cannot forecast two channels with transfer entropy theory.

We will show the test results in section VI Simulations and Performance Analysis.

III. 5G CHANNEL FORECASTING BASED ON TRANSFER ENTROPY

A. INTRODUCTION TO TRANSFER ENTROPY

In this paper, we found that the 5G channel coefficient followed the Gaussian distribution. Based on the equivalence between Granger causality and transfer entropy under Gaussian variables [16], we could use the transfer entropy to forecast the 5G channel coefficient.

Transfer Entropy is a time-asymmetry non-parametric information measure based on conditional co-information proposed by Schreiber [17]. Although it differs from Granger causality based on the vector auto-regression model, transfer entropy does not use any model assumptions. Both are essentially derived from Wiener's construction of causality, that is, the addition of historical information of new variables, which reduces the uncertainty of the prediction of the target variable [2], [18]. And causality is a measure of this change in uncertainty.

The definition of transfer bribery: For two discrete random variables X and Y, the probability distribution functions are $p(x)$ and $p(y)$ respectively. The joint probability of events x and y occurring at the same time is p(x,y), then the Shannon entropy H(x) is defined as [19].

$$H(x) = -\sum_{x} p(x) log p(x) \tag{5}$$

The conditional probability of Y with X is:

$$p(y|x) = p(x, y)/p(x) \tag{6}$$

In Eq.(6) if $p(y|x) = p(y)$ x and y are independent. Then $p(x, y) = p(y|x)p(x) = p(y)p(x)$

The joint entropy of X and Y is

$$H(X, Y) = -\sum_{x,y} p(x, y) log p(x, y) \tag{7}$$

The conditional entropy of X with Y is

$$H(X|Y) = -\sum_{x,y} p(x, y) log p(x|y) \tag{8}$$

The Mutual Information(MI) contained between X and Y is the output of the two systems as though they were independent as opposed to their 'actual' relationship

$$\begin{aligned} &M(X; Y) \\ &\quad = H(X) + H(Y) - H(X, Y) \\ &\quad = -\sum_{x,y} p(x, y) log(p(x)p(y)) + \sum_{x,y} p(x, y) log(p(x, y)) \\ &\quad = \sum_{x,y} p(x, y)[log(p(x, y)) - log(p(x)p(y))] \\ &\quad = \sum_{x,y} p(x, y) log \frac{p(x, y)}{p(x)p(y)} \end{aligned} \tag{9}$$

However, MI is not effective at predicting future events from current data since it is symmetric, $M(X, Y) = M(Y, X)$. And it does not indicate which way the information is flowing. These shortcomings may be remedied by time shifting one of the variables. Transfer Entropy (TE) (Schreiber 2000 [17]) is based on rates of entropy change, it captures some of the dynamics of a system.

Suppose two systems which generate events. We define an entropy rate which is the amount of additional information required to represent the value of the next observation of one of the systems [20]:

$$h_1 = -\sum_{x_{n+1}} p(x_{n+1}, x_n, y_n) log_a p(x_{n+1}|x_n, y_n) \tag{10}$$

Suppose that value of observation x_{n+1} was not dependent on the current observation y_n:

$$h_2 = -\sum_{x_{n+1}} p(x_{n+1}, x_n, y_n) log_a p(x_{n+1}|x_n) \tag{11}$$

The quantity h_1 represents the entropy rate for the two systems, and h_2 represents the entropy rate assuming that x_{n+1} is independent of y_n.Thus, we get transfer entropy:

$$\begin{aligned} h_2 - h_1 = &- \sum_{x_{n+1}, x_n, y_n} p(x_{n+1}, x_n, y_n) log_a p(x_{n+1}|x_n) \\ &+ \sum_{xn+1, x_n, y_n} p(x_{n+1}, x_n, y_n) log_a p(x_{n+1}|x_n, y_n) \\ = &\sum_{x_{n+1}, x_n, y_n} p(x_{n+1}, x_n, y_n) log_a \left(\frac{p(x_{n+1}|x_n, y_n)}{p(x_{n+1}|x_n)}\right) \end{aligned} \tag{12}$$

There are actually two equations for the transfer entropy, because it has an inherent asymmetry in it.

$$T_{J\to I} = \sum_{x_{n+1}, x_n, y_n} p(x_{n+1}, x_n, y_n) log_a \left(\frac{p(x_{n+1}|x_n, y_n)}{p(x_{n+1}|x_n)}\right) \tag{13}$$

$$T_{I\to J} = \sum_{y_{n+1}, x_n, y_n} p(y_{n+1}, x_n, y_n) log_a \left(\frac{p(y_{n+1}|x_n, y_n)}{p(y_{n+1}|y_n)}\right) \tag{14}$$

Then with substitutions:

$$\begin{aligned} p(x_{n+1}|x_n, y_n) &= p(x_{n+1}, x_n, y_n)/p(x_n, y_n) \\ p(x_{n+1}|x_n) &= p(x_{n+1}, x_n)/p(x_n) \end{aligned}$$

our equations become [20]

$$T_{J\to I} = \sum_{x_{n+1}, x_n, y_n} p(x_{n+1}, x_n, y_n) \cdot log \left(\frac{p(x_{n+1}, x_n, y_n) \cdot p(x_n)}{p(x_n, y_n) \cdot p(x_{n+1}, x_n)}\right) \tag{15}$$

$$T_{I\to J} = \sum_{y_{n+1}, x_n, y_n} p(y_{n+1}, x_n, y_n) \cdot log \left(\frac{p(y_{n+1}, x_n, y_n) \cdot p(y_n)}{p(x_n, y_n) \cdot p(y_{n+1}, y_n)}\right) \tag{16}$$

In (15)(16), J represents 5G channel 1, and I represents 5G channel 2. We use TE algorithm to forecast the channels based on two equations. Since the row 5G channels are too long to plot, we choose part of the forecasted channel randomly, then show them in section VI Simulations and Performance Analysis.

IV. FORECASTED 5G CHANNEL ACCURACY PROOF BASED ON RMSE AND CRLB

A. ROOT MEAN SQURE ERROR OF FORECASTED CHANNEL

First we apply the RMSE metric between real 5G channel and forecasted channel. The RMSE can be computed as

$$RMSE = \sqrt{\sum_{n=1}^{N}((x_n - x_n')^2)/N} \tag{17}$$

where x_n is the observed value, which is the forecasted channel coefficient in our work. x_n' is the true value, which represents the real 5G channel coefficients here. N is the total number of channel coefficient. In order to avoid specialty, we proposed 2000 pairs of Real 5G channels as training groups and 1000 pairs of Real 5G channels as test groups to do the forecasting based on transfer entropy and calculated the average RMSE. We will show the RMSE performance in section VI Simulations and Performance Analysis.

B. CRAMER-RAO LOWER BOUND

Cramer-Rao Lower Bound (CRLB) can be used to calculate the best estimation accuracy that can be obtained in unbiased estimation, so it is often used to calculate the best estimation accuracy that can be achieved by theoretically, and to evaluate the performance of parameter estimation methods (Whether it is close to the lower limit of CRLB) [21].

The 5G channel follows a Gaussian distribution with a variance γ^2. Assume $x = A + \omega$, ω is a Gaussian noise, $\omega \sim N(0, \sigma^2)$.

The pdf of x_m is [22]

$$f(x_m) = \frac{1}{\sqrt{2\pi(\gamma^2+\sigma^2)}} exp\left[-\frac{(x_m-m)^2}{2(\gamma^2+\sigma^2)}\right] \tag{18}$$

where x_m is a random channel coefficient, σ^2 is the variance of Gaussian noise ω.

Let $x \triangleq [x_1, x_2, \ldots, x_M]$, then the pdf of x is

$$\begin{aligned} f(x) &= \prod_{m=1}^{M} f(x_m) \\ &= \prod_{m=1}^{M} \frac{1}{\sqrt{2\pi(\gamma^2+\sigma^2)}} exp\left[-\frac{(x_m-m)^2}{2(\gamma^2+\sigma^2)}\right] \end{aligned} \tag{19}$$

let

$$\theta \triangleq \gamma^2$$

then Eq. (19) can be expressed as

$$f(x) = \prod_{m=1}^{M} \frac{1}{\sqrt{2\pi(\theta+\sigma^2)}} exp\left[-\frac{(x_m-m)^2}{2(\theta+\sigma^2}\right] \tag{20}$$

Then find the logarithm to get the log likelihood function

$$logf(x) = \sum_{m=1}^{M}\left[log\frac{1}{\sqrt{2\pi(\theta+\sigma^2)}}\right] + \sum_{m=1}^{M}\frac{-(x-m)^2}{2(\theta+\sigma^2)} \tag{21}$$

let

$$\frac{\partial}{\partial\theta} logf(x)|_{\theta=\hat{\theta}} = \sum_{m=1}^{M}\left[\frac{-1}{2(\theta+\sigma^2)} + \frac{(x-m)^2}{2(\theta+\sigma^2)^2}\right] = 0 \tag{22}$$

which has the unique solution

$$\hat{\theta}(x) = \sum_{m=1}^{M} \sigma^2 - (x-m)^2 \tag{23}$$

Since

$$\frac{\partial^2}{\partial\theta^2} logf(x)|_{\theta=\hat{\theta}} = \sum_{m=1}^{M}\frac{-1}{2(\theta+\sigma^2)} - \frac{(x-m)^2}{(\theta+\sigma^2)^3} < 0 \tag{24}$$

this solution gives the unique maximum of $logf(x)$. The expectation of $\hat{\theta}(x)$ is

$$\begin{aligned} E\left[\hat{\theta}(x)\right] &= \int_0^{\infty}[\sum_{m=1}^{M}\sigma^2 - (x-m)^2]f(x_m)d_{x_m} \\ &= \int_0^{\infty}\left[\sum_{m=1}^{M}\sigma^2-(x-m)^2\right] \\ &\quad \cdot \frac{1}{\sqrt{2\pi(\theta+\sigma^2)}} exp\left[-\frac{(x_m-m)^2}{2(\theta+\sigma^2)}\right] d_{x_m} \\ &= \theta \end{aligned} \tag{25}$$

Therefore it's unbiased. The Fisher's information for our case can be expressed as

$$\begin{aligned} I_\theta &= -E_\theta\left[\frac{\partial^2}{\partial\theta^2} logf(x)\right] \\ &= -E_\theta\left[\sum_{m=1}^{M}\frac{-1}{2(\theta+\sigma^2)} - \frac{(x-m)^2}{(\theta+\sigma^2)^3}\right] \end{aligned} \tag{26}$$

So the Cramer-Rao lower bound (CRLB) is

$$Var_\theta[\hat{\theta}(x)] \geq \frac{1}{I_\theta} = \frac{1}{-E_\theta\left[\sum_{m=1}^{M}\frac{-1}{2(\theta+\sigma^2)} - \frac{(x-m)^2}{(\theta+\sigma^2)^3}\right]} \tag{27}$$

We will show the CRLB performance in section VI Simulations and Performance Analysis.

V. CHANNEL POWER ALLOCATION BASED ON INVERSE WATER-FILLING ALGORITHM

The Inverse Water-Filling(IWF) algorithm is based on a certain criterion, and adaptively allocates the transmission power according to the channel condition. Usually when the channel condition is good, then the power is allocated more, and when the channel is poor, the power is allocated less, and when the channel is poor, the power is not allocated. Thereby maximizing the transmission rate. To achieve the 'IWF' distribution of power, the transmitter must know the Channel State Information(CSI). When the transmitter knows the channel, the channel capacity can be increased [23].

Consider a $r \times 1$-dimensional zero-mean cyclic symmetric complex Gaussian signal vector $\tilde{s}$, where r is the rank of the transmission channel. The vector is multiplied by the matrix $V(H = U\Sigma V^H)$ before transmission. At the receiver, the received signal vector y is multiplied by U^H [24]. The effective input-output relationship for this system is given by:

$$\tilde{y} = \sqrt{\frac{E_s}{M_T}} U^H H V \tilde{s} + U^H n = \sqrt{\frac{E_s}{M_T}} U^H U \Sigma V^H \tilde{s} + U^H n$$
$$= \sqrt{\frac{E_s}{M_T}} \Sigma \tilde{s} + \tilde{n} \tag{28}$$

where $\tilde{y}$ is the received signal vector of the $r \times 1$ dimensional transform, and $\tilde{n}$ is the zero-mean cyclic symmetric complex Gaussian $r \times 1$ transform noise vector whose covariance matrix is $\xi\{\tilde{n}\tilde{n}_H\} = N_0 I_r$. The vector $\tilde{s}$ must satisfy $\xi\{\tilde{s}\tilde{s}_H\} = M_T$ to limit the total transmitted energy.

Then

$$\tilde{y}_i = \sqrt{\frac{E_s}{M_T}} \sqrt{\lambda_i} \tilde{s}_i + \tilde{n}_i, \quad i = 1, 2, \ldots, r \tag{29}$$

The channel capacity is given by following

$$C = \sum_{i=1}^{r} log_2(1 + \frac{E_s \gamma_i}{M_T N_0} \lambda_i) \tag{30}$$

where $\gamma_i = \xi\{|s_i|^2\}(i = 1, 2, \ldots, r))$ reflects the transmission energy of the i-th subchannel and satisfies $\sum_{i=1}^{r} \gamma_i = M_T$, E_s represents the channel gain on the s_{th} subchannel, λ_i is the Lagrange multiplier.

Variable energy can be allocated in the subchannel to maximize mutual information. Now the problem of maximizing energy becomes

$$C = \max_{\sum_{i=1}^{r}} y_i = M_T \sum_{i=1}^{r} log_2(1 + \frac{E_s \gamma_i}{M_T N_0} \lambda_i) \tag{31}$$

Maximizing by Lagrangian method. The Optimal energy allocation policy is

$$\gamma_i^{opt} = max\{(\mu - \frac{M_T N_0}{E_s \lambda_i}), 0\} \tag{32}$$

where

$$\sum_{i=1}^{r} \gamma_i^{opt} = M_T \tag{33}$$

We compared the channel capacity based on IWF algorithm and the equal gain algorithm. We will show the simulations in section VI Simulations and Performance Analysis.

VI. SIMULATIONS AND PERFORMANCE ANALYSIS

A. SIMULATION PARAMETERS OF SIMULATED CHANNELS

Table 1 shows the simulation parameters we used for simulated channels.

TABLE 1. Simulation parameters.

Frequency (GHz)	28.0	Bandwidth (MHz)	800
TXPower (dBm)	30.0	Environment	NLOS
Scenario	UMi	Pressure (mBar)	1013.25
Humidity	50	Temperature (Celsius)	20.0
RainRate (mm/hr)	0.0	Polarization	Co-Pol
Foliage	No	DistFol (m)	0.0
FoliageAttenuation (dB)	0	TxArrayType	ULA
RxArrayType	ULA	Num of TXElements	1
Nub=m of RXElements	1	TXAziHPBW	10
TXElvHPBW	10	RXAziHPBW	10
RXElvHPBW	10		

B. GRANGER CAUSALITY TEST OF TWO RANDOM 5G CHANNELS

Table 2 shows the results of the Granger causality tests.

TABLE 2. Granger causality test results.

	F-Value	Critical Value	Confidence Level
1 on 2	20.9127	0.229	0.95
2 on 1	97.0205	0.229	0.95

First, we do a Granger causality test of channel 1 on channel 2. In this simulation, the value of the F statistic is 20.9127, where the critical value from the F-distribution is 0.229, confidence level α is 0.95. Since the value of the F statistic is larger than the critical value from the F-distribution, we reject the null hypothesis that channel 2 does not Granger Cause channel 1. So this test proves that channel 2 is Granger cause of channel 1. Then we do the Granger causality test of channel 2 on channel 1. In this case, the value of the F statistic is 97.0205, where the critical value from the F-distribution is still 0.229, confidence level α is 0.95. So this test proves that channel 2 is also the Granger cause of channel 1.

So far, we verified the existence of Granger causality between the two channels. Based on this result, we use transfer entropy to forecast the two channels.

C. FORECASTED CHANNEL BASED ON TRANSFER ENTROPY

In Fig.1 and Fig.2, we compare the real 5G channel coefficient with the predicted 5G channel coefficient. Observe that the channels which are forecasted using the TE algorithm has high accuracy. But we still need a mathematical method to calculate the specific error between the two channels and prove the accuracy. Therefore, for the comparison of real

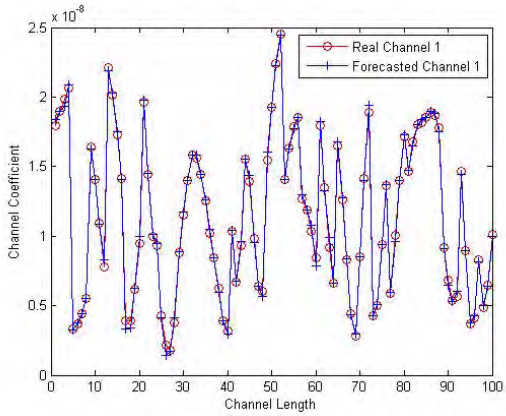


FIGURE 1. Forecasted 5G channel 1 based on transfer entropy.

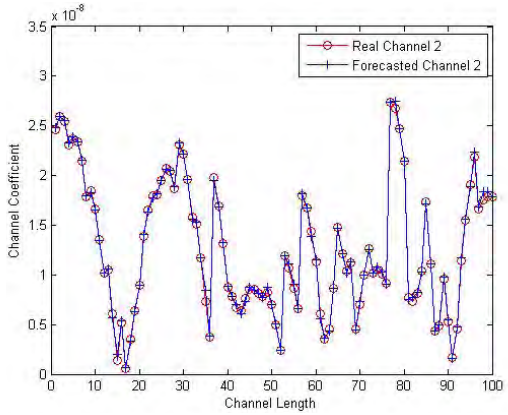


FIGURE 2. Forecasted 5G channel 2 based on transfer entropy.

channels and prediction channels, we apply two metrics, i.e., RMSE and CRLB.

D. RMSE

In Fig. 3, we purposed the Box Jenkin's method and the Transfer Entropy method to forecast the real channels and compared their RMSE, which is the average RMSE of all 1000 pairs forecasted channels and 1000 pairs real 5G channels with different SNRs.

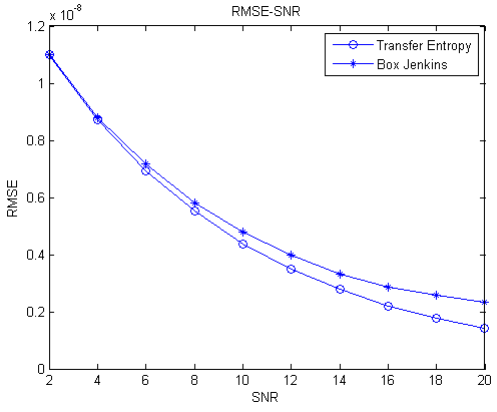


FIGURE 3. The average RMSE between the real 5G channel and forecasted channel.

1) We can observe from Fig. 1 and Fig. 2 that the average coefficient of the real 5G channel is about 1.5×10^{-8}. And in Fig. 3, the RMSE is really high when the SNR is low. So we can conclude that we compare the TE method and Box Jenkin's method when the SNR is low, the accuracy is insufficient.

2) However, with the SNR is increasing, the RMSE is going to horizontally and stable at about 1×10^{-9}. So this shows that the true RMSE of the real 5G channel and forecasted channel based on the TE algorithm is less than 6%. This result proves that TE is feasible to forecast 5G channels.

3) From Figure 3 we can observe that under the same SNR, the RMSE between the forecasted channel and real channel by using the Box Jenkins method is larger than that of the Transfer Entropy method, which means that the TE method is more accurate in the forecasting under the same channel status.

E. CRAMER-RAO LOWER BOUND

Simulations: For real and forecasted 5G channel1 and channel 2, we run Monte Carlo simulations at each SNR value and applied $\hat{\theta}$ for total 4 channel respectively. In Fig. 4 and Fig. 5, we plotted the variance of the estimator with different values of SNRs. Observe the following.

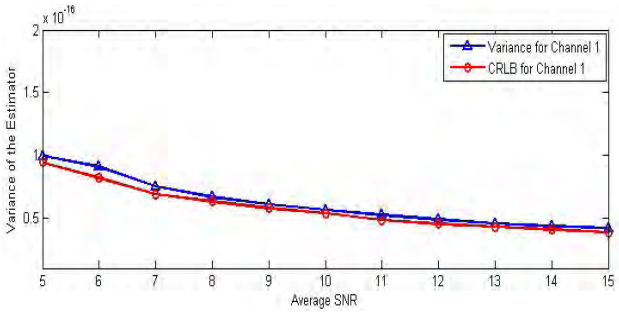


FIGURE 4. Variance of forecasted channel 1 with different value of SNRs.

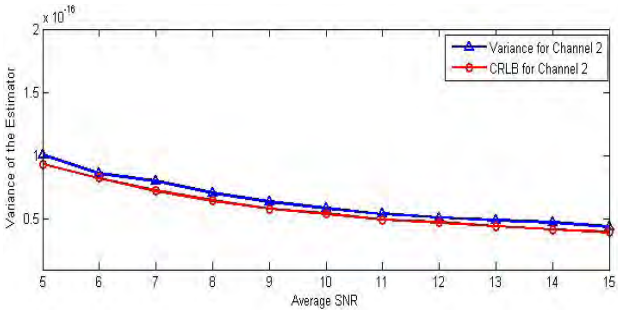


FIGURE 5. Variance of forecasted channel 2 with different value of SNRs.

1) The actual variance of $\hat{\theta}$ is almost mach with the CRLB for different SNR value [25], which validate our result in section A: Our forecasted 5G channel is close to real 5G channel and TE is feasible to forecast 5G channels.

2) The actual variance of $\hat{\theta}$ reduces as SNR value increases, and tends to stable and horizontally, which is as we have shown the RMSE plots in section A.

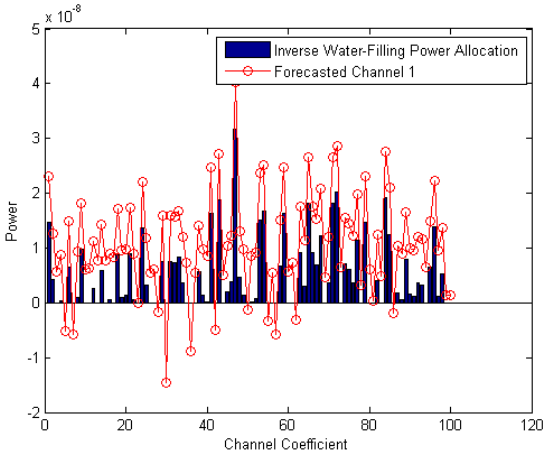


FIGURE 6. Forecasted 5G channel 1 power allocation based on inverse water-filling.

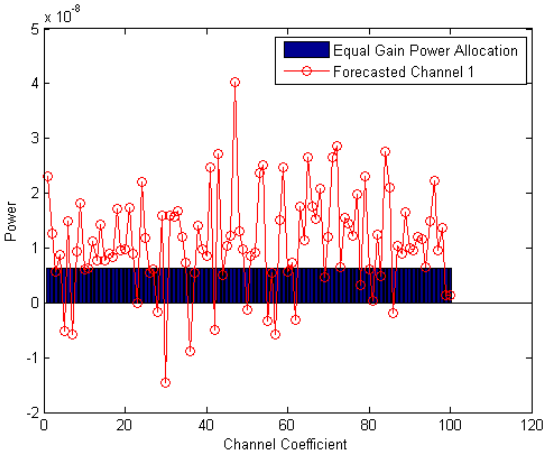


FIGURE 7. Forecasted 5G channel 1 power allocation based on equal gain.

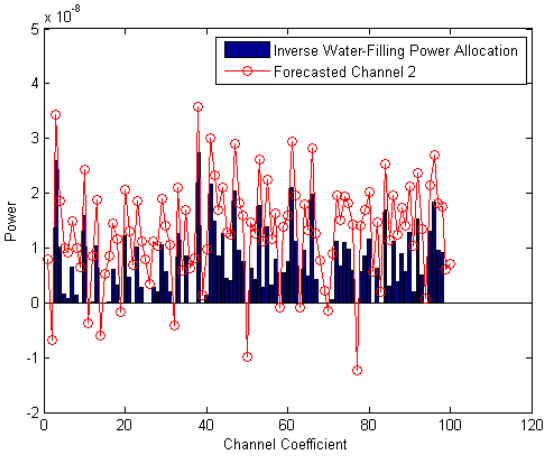


FIGURE 8. Forecasted 5G channel 2 power allocation based on inverse water-filling.

F. CHANNEL POWER ALLOCATION

Simulations: We perform power allocation based on the IWF algorithm and Equal Gain(EG) algorithm for the forecasted

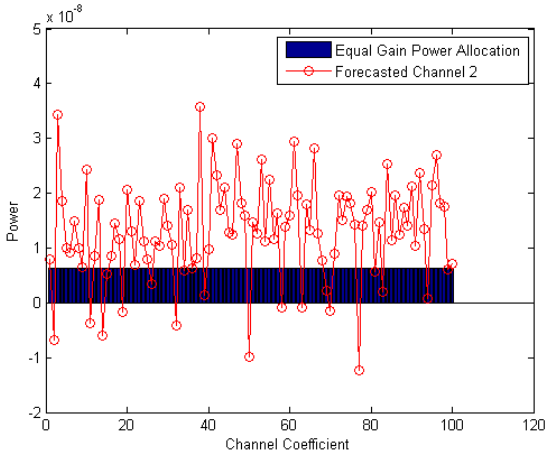


FIGURE 9. Forecasted 5G channel 2 power allocation based on equal gain.

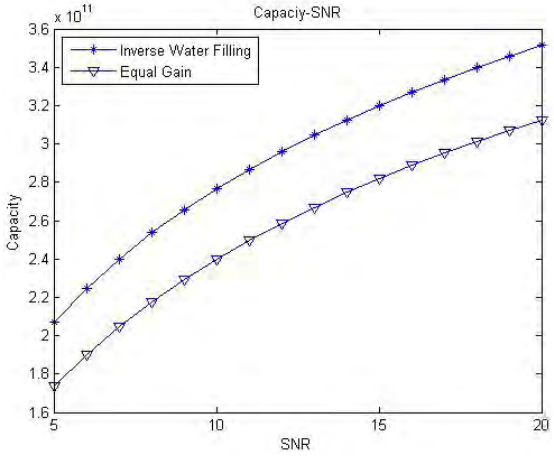


FIGURE 10. Average forecasted channel capacity comparison based on inverse water-filling and equal gain.

5G channel 1 and channel 2 and plot part of the results respectively. Since the row 5G channels are too long to plot, we choose part of the forecasted channel randomly, so there are some variations in the channel coefficients compared with transfer entropy. Fig. 6- Fig. 9 shows the simulations. For each simulation, the total power is $1 \times 10^{-5}w$, which is $-20dBm$ to be allocated in the channel. And Fig. 10 proposed the channel capacity comparison. Observe the following.

1) Fig. 6 and 8 shows power allocation using the IWF algorithm. It can be seen from the figures that when the subchannel state is good, the much power is allocated to the subchannel, and when the subchannel is not good, the less power is allocated. Some subchannels are very poor, and no power is allocated to the channel. Compared with the power allocation of the EG algorithm in Fig. 7 and Fig. 9, that is, the energy is evenly distributed on each subchannel regardless of the channel condition, and the IWF algorithm can greatly improve the energy efficiency and avoid the waste of power.

2) From Figure 10, we can observe that at the same SNR, the channel capacity of the IWF is greater than the EG, where demonstrates the power allocation efficiency of the IWF algorithm working on the forecasted 5G channels.

VII. CONCLUSIONS AND FUTURE WORKS

In this paper, we have proposed the channel forecasting in 5G Wireless Communication Based on Granger Causality and Transfer Entropy, and the power allocation based on Inverse Water-filling algorithm. Our work consists of four main parts: Firstly, we performed the Granger causality test on two independent, random and real 5G channels of the simulation, ensuring the correlation between the two channels. Secondly, we used the transfer entropy method to forecast the above two 5G channels and obtain two forecasted 5G channels. Third, we calculated the RMSE of the original channel and the forecasted channel to ensure the high accuracy of the forecasting. At the same time, we computed the CRLB of the forecasted 5G channel and showed that the variance of the forecasted parameters is close to the CRLB. Finally, for the two forecasted 5G channels, we performed power allocation and comparison based on the Equal Gain algorithm and the Inverse Water-filling algorithm. Simulations further validate these theoretical results.

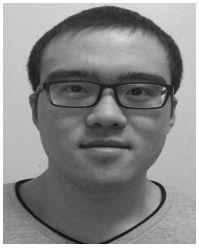
The long-term concern has been to provide high throughput and data transfer rates, but now we have realized that high energy efficiency is equally important even with little or no data transmission and processing. So in the future, we will try to find more algorithms to achieve the power allocation efficiency of 5G channels.

ACKNOWLEDGMENTS

The authors would like to thank the New York University NYU WIRELESS center for providing an open source 5G channel simulator used for simulating 5G channel data.

REFERENCES

[1] Z. Wang, Z. Luo, and K. Wei, "5G service requirements and progress on technical standards," *ZTE Technol. J.*, vol. 20, no. 2, pp. 1–4, 2014.
[2] N. A. Ahmed and D. V. Gokhale, "Entropy expressions and their estimators for multivariate distributions," *IEEE Trans. Inf. Theory*, vol. 35, no. 3, pp. 688–692, May 1989.
[3] R. Baldemair *et al.*, "Future wireless communications," in *Proc. IEEE 77th Veh. Technol. Conf. (VTC Spring)*, Jun. 2013, pp. 1–5.
[4] S. Liu, J. Wu, C. H. Koh, and V. K. N. Lau, "A 25 Gb/s(/km^2) urban wireless network beyond IMT-advanced," *IEEE Commun. Mag.*, vol. 49, no. 2, pp. 122–129, Feb. 2011.
[5] A. Shojaie and G. Michailidis, "Discovering graphical Granger causality using the truncating lasso penalty," *Bioinformatics*, vol. 26, no. 18, pp. i517–i523, Sep. 2010.
[6] Y. Park, "5G vision and requirements of 5G forum," in *5G Forum*. 2014.
[7] R. Mumford, "Anite leads 5G radio channel model development," Horizon House, Norwood, MA, USA, Tech. Rep., 2013.
[8] A. Osseiran *et al.*, "Scenarios for 5G mobile and wireless communications: The vision of the METIS project," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 26–35, May 2014.
[9] X. H. You, Z. W. Pan, and X. Q. Gao, "The 5G mobile communication: The development trends and its emerging key techniques," *Sci. China*, vol. 44, no. 5, pp. 551–563, 2014.
[10] I. Hwang, B. Song, and S. S. Soliman, "A holistic view on hyper-dense heterogeneous and small cell networks," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 20–27, Jun. 2013.
[11] C.-X. Wang *et al.*, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 122–130, Feb. 2014.
[12] C. W. J. Granger, "Testing for causality: A personal viewpoint," *J. Econ. Dyn. Control*, vol. 2, no. 1, pp. 329–352, 1980.
[13] F. Han, H. Lu, and H. Liu, "A direct estimation of high dimensional stationary vector autoregressions," *J. Mach. Learn. Res.*, vol. 16, no. 1, pp. 3115–3150, Jan. 2015.
[14] K. Zhou, H. Zha, and L. Song, "Learning social infectivity in sparse low-rank networks using multi-dimensional Hawkes processes," in *Proc. JMLR Workshop Conf.*, 2013, pp. 641–649.
[15] E. C. Hall, G. Raskutti, and R. Willett. (May 2016). "Inference of high-dimensional autoregressive generalized linear models." [Online]. Available: https://arxiv.org/abs/1605.02693
[16] L. Barnett, A. B. Barrett, and A. K. Seth, "Granger causality and transfer entropy are equivalent for Gaussian variables," *Phys. Rev. Lett.*, vol. 103, no. 23, Dec. 2009, Art. no. 238701.
[17] T. Schreiber, "Measuring information transfer," *Phys. Rev. Lett.*, vol. 85, no. 2, p. 461, 2000.
[18] H. Qiu, S. Xu, F. Han, H. Liu, and B. Caffo, "Robust estimation of transition matrices in high dimensional heavy-tailed vector autoregressive processes," in *Proc. Int. Conf. Mach. Learn.*, vol. 37, 2015, p. 1843.
[19] G. Aulogiaris and K. Zografos, "A maximum entropy characterization of symmetric Kotz type and burr multivariate distributions," *Test*, vol. 13, no. 1, pp. 65–83, Jun. 2004.
[20] M. Bicego, D. Gonzalez-Jimenez, E. Grosso, and J. L. A. Castro, "Generalized Gaussian distributions for sequential data classification," in *Proc. 19th Int. Conf. Pattern Recognit.*, Dec. 2008, pp. 1–4.
[21] K. Zografos and S. Nadarajah, "Expressions for Rényi and Shannon entropies for multivariate distributions," *Statist. Probab. Lett.*, vol. 71, no. 1, pp. 71–84, Jan. 2005.
[22] I. S. Merrill *et al.*, *Introduction to Radar Systems*, vol. 7, no. 10. New York, NY, USA: McGraw-Hill, 2001.
[23] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications, and Control*. London, U.K.: Pearson, 1995.
[24] R. D. Palmer, "Fundamentals of radar signal processing," *Bull. Amer. Meteorological Soc.*, vol. 89, no. 7, p. 1037, 2008.
[25] M. Zhong, F. Lotte, M. Girolami, and A. Lécuyer, "Classifying EEG for brain computer interfaces using Gaussian processes," *Pattern Recognit. Lett.*, vol. 29, no. 3, pp. 354–359, Feb. 2008.



ZHANGLIANG CHEN received the B.S. degree from Tianjin University, Tianjin, China, in 2016. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, The University of Texas at Arlington, Arlington, TX, USA. His current research interests include wireless communications, massive MIMO, radar sensor networks, wireless sensor networks, and signal processing.

QILIAN LIANG (A'01–M'01–SM'05–F'17) received the B.S. degree in electrical engineering from Wuhan University, Wuhan, China, in 1993, the M.S. degree in electrical engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 1996, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 2000.

He was a member of Technical Staff with Hughes Network Systems Inc., San Diego, CA, USA. He is a Distinguished University Professor with the Department of Electrical Engineering, The University of Texas at Arlington (UTA), Arlington, TX, USA. He has authored or coauthored over 300 journals and conference papers, seven book chapters, and has six U.S. patents pending. His current research interests include radar sensor networks, wireless sensor networks, wireless communications, compressive sensing, smart grids, signal processing for communications, and fuzzy logic systems and applications.

Dr. Liang was a recipient of the 2002 IEEE Transactions on Fuzzy Systems Outstanding Paper Award, the 2003 U.S. Office of Naval Research Young Investigator Award, the 2005 UTA College of Engineering Outstanding Young Faculty Award, the 2007, 2009, and 2010 U.S. Air Force Summer Faculty Fellowship Program Award, the 2012 UTA College of Engineering Excellence in Research Award, and the 2013 UTA Outstanding Research Achievement Award. He was inducted into the UTA Academy of Distinguished Scholars, in 2015.