Data-Driven Channel Modeling Using Spectrum Measurement

Shang-Pin Sheng, Mingyan Liu, *Fellow, IEEE*, and Romesh Saigal

Abstract—Dynamic spectrum access has been a subject of extensive study in recent years. The increasing volume of literature calls for better understanding of the characteristics of current spectrum utilization as well as better tools for analysis. A number of measurement studies have been conducted recently, revealing previously unknown features. On the other hand, analytical studies largely continues to rely on standard models like the two-state Markov (Gilbert-Elliot) model. In this paper, we present an alternative, stochastic differential equation (SDE) based spectrum utilization model that captures dynamic changes in channel conditions induced by primary users' activities. The SDE model is in closed form, can generate spectrum dynamics as a temporal process, and is shown to provides very good fit for real spectrum measurement data. We show how synthetic spectrum data can be generated in a straightforward manner using this model to enable realistic simulation studies. Moreover, we show that the SDE model can be viewed as a more general modeling framework (continuous in time and continuous in value) than commonly used discrete Markovian models: it is defined by only a few parameters but can be used to obtain the transition matrix of any N-state Markov model. This is verified by comparing the two-state GE model generated by the SDE model and that trained directly from the data. We show that the GE model is a good fit for the (quantized) data, thereby a fine choice when binary descriptions of the channel condition is sufficient. However, when highly resolution (in channel condition) is needed, the SDE model is much more accurate than an N-state model, and is much easier to train and store.

Q1

1 Introduction

Communications over wireless media is generally much more error prone compared to their wireline counterparts due to the noisier, time-varying and often unpredictable nature of the wireless channel quality. Consequently, the modeling of error patterns in wireless communication has been an important tool for evaluating the performance of wireless communication and networking algorithms, both in analysis and in simulation. Research in recent years on dynamic (and open) spectrum access using cognitive radios (CR) [1] has further exemplified the importance of channel modeling: the ability of wireless devices to quickly detect spectrum availability and exploit instantaneous spectrum opportunities is greatly enhanced if they are equipped with a good channel model that accurately captures the stochastic behavior of the underlying channel conditions. This is reflected in our ability to derive better channel sensing and access decisions both in theory and in numerical experiment when good channels models are available.

1.1 Related Work on Channel Modeling

Within this application context, the desire to better understand spectrum utilization, especially in the presence of licensed primary users, has motivated a series of spectrum measurement studies published recently, see e.g., [2], [3], [4], [5]. These measurement studies, however, have not in general led to tools that can generate realistic spectrum utilization as a time process to evaluate spectrum sensing and access algorithms. In [6] a sequence of probability distributions of spectrum availability were derived using measurement data. However, these distributions capture only the average behavior of spectrum rather than describing spectrum activity as a process in time as we aim to achieve in this paper. Motivated by this, our goal in this paper is to construct stochastic models that can capture key properties of wireless channels that are important in evaluating opportunistic spectrum access schemes.

One commonly used category of channel models is based on Markov chains, where each state often represents a different condition of the channel, with dynamic changes described by the state transition probability matrix. Sometimes a mixture of Markov models with other models is used to capture characteristics of the error patterns, see e.g., the chaotic maps [7] and the MTA [8].

Within this category, the Gilber-Elliott (GE) model [9], [10] is the simplest Markov model consisting of only two states. Perhaps due to its simplicity (and often the associated analytical tractability), the GE model is widely used as the underlying model for wireless channels both in analysis and in simulation. Under this model, the channel is given by a two-state Markov chain with state G (good) and B (bad), see Fig. 8. In state G, transmission is assumed error-free, while in state B the channel has a probability h of transmitting the packet correctly.[1] These

- *S.-P. Sheng and M. Liu are with the Electrical Engineering and Computer Science Department, University of Michigan, Ann Arbor, Michigan. E-mail: {shangpin, mingyan}@umich.edu.*
- *R. Saigal is with the Industrial and Operations Engineering Department, University of Michigan, Ann Arbor, Michigan. E-mail: rsaigal@umich.edu.*

Q2

Manuscript received 15 Nov. 2013; revised 3 Sept. 2014; accepted 3 Nov. 2014. Date of publication 0 . 0000; date of current version 0 . 0000.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TMC.2014.2374152

1. The GE model is very often used in a simplified version where $h = 0$.

two states are used to model a burst-noise channel. A more general variation of this model includes a probability k (usually $k > h$) such that both good and bad states have a chance to generate an error bit.

Many studies have been conducted on these two-state Markov models. McDougall and Miller [11] showed that at low SNR, the two-state Markov model does not generate an adequate frame error process because it lacks the ability to match higher-order block error statistics. Hartwell and Fapojuwo [12] showed that using the higher-order state hidden Markov models provides a better fit of measured data than the traditional two-state GE models. However, high-order state Markov models require high computational complexity to train the parameters of the Markov model. Yu and Miller [13] proposed a four-state Markov model and showed how to analytically establish the transition probability. Konrad et al. [8] proposed a Markov based model aimed at capturing the non-stationary behaviors of wireless channels.

1.2 Our Approach and Main Contributions

We introduce a stochastic differential equation (SDE) model derived partly based on the physics of electromagnetic wave propagation. This SDE model is continuous both in time and in value, and falls under the category of *diffusion models*, which is more commonly used in queuing analysis when dealing with large systems, e.g., those with heavy loads [14]. The main idea is that when queue sizes are large, the increments over a single discrete step become relatively small by comparison. Thus under such operating regimes, it is reasonable to model the discrete change in the queue occupancy by a continuous flow, resulting in diffusion models. The analytical advantage of using a diffusion model is that it is amenable to both transient and steady state analysis and can be used to derive queue size distributions which is hard to do using a discrete model when the state space is large. We shall see that the SDE model introduced here holds similar advantages over discrete, GE-type models. We therefore conclude that the SDE model serves as a valuable alternative to the commonly used GE-type models.

Our main results are summarized as follows.

1) We use real data collected from spectrum measurement studies to verify the SDE model and it is shown to fit the data very well.
2) We show by using this model how to synthesize sample paths (temporal power process) of a wireless channel, thereby creating a realistic spectrum environment which can be used for simulation studies. We compare the spectral entropy of the synthesized sample paths with the collected data; results show that they are consistent.
3) The proposed SDE model can be used to generate the two-state Markov (GE) model (and also can generate an N-state Markov models) through time-discretization and value-quantization. We show that the resulting GE model is accurate in representing the quantized version of the original data, in terms of the run length distribution and power spectral density. This SDE-generated GE model also matches closely the GE model generated directly from the quantized data. Therefore the SDE model may be viewed as a continuous generalization of the discrete GE model (and more broadly a discrete N-state model), and while the former can be used to obtain the latter the reverse is not true.
4) We compare the SDE with an N-state Markov model (both directly generated by the same set of data) by using them to predict channel conditions. Our results show that to achieve the same level of prediction accuracy, the N-state Markov model requires much higher complexity in computation and training due to the need to minimize the quantization error and obtain accurate transition probabilities. We further show that if we obtain the transition matrix of the N-state Markov model from the SDE model, then the error introduced in the training stage can be eliminated. However, it is still much easier to predict using the SDE model directly at the same level of accuracy.
5) Last but not least, it is worth noting that the SDE model is defined by only three parameters and is thus very easy and inexpensive to train with much less data compared to a discrete N-state model. It is also much more robust to imperfections in the data, e.g., when samples are not exactly collected at constant intervals. Furthermore, once the SDE model is trained, it can be used to at any desired time resolution due to its continuous-time nature, whereas an N-state Markov model would need to be retrained if one wants to reduce the size of the discrete time step (i.e., increase the time resolution).

We emphasize that the SDE model is intended as a generic channel model, and in particular as an alternative to the commonly used GE model; its applicability is therefore quite broad.

1.3 Organization

The remainder of the paper is organized as follows. We introduce the SDE model and show how to estimate its parameters in Section 2. In Section 3 we verify the model by using spectrum measurement data from CRAWDAD [15] as well as our own study [5]. In Section 2.4 we show how to synthesize data from the trained SDE model and compare the spectral entropy of the synthesized data to the spectral entropy of the collected data. Then in Section 4 we show that we can obtain the two-state GE model from the synthesized data and compare it with the GE model trained directly from the quantized measurement data. In Section 5 we compare the channel prediction performance of the SDE model and an N-state Markov model, both trained from the measurement data. We conclude the paper in Section 6.

2 The SDE Model

2.1 Constructing the Model

The spectrum utilization model presented here uses stochastic differential equations to model dynamic scattering and multipath fading channels, in particular, Rayleigh-distributed stationary channels. This is a technique developed and used in a number of studies, see e.g., [16], [17].

Specifically, our model is derived from a dynamic wireless channel model developed in [18] using similar techniques. Underlying this model is the assumption of either a single transmitter or many non-dominant transmitters stationary in space and in time. The model describes the complex signal received by a stationary receiver (thus with zero Doppler's effect) Building upon this work, our contribution lies in (1) extracting the received energy as a random process expressed as an SDE and the construction of the subsequent spectrum model, and (2) developing a method to estimate the unknown parameters of the model.

In this model the signal detected at a receiver is viewed as a collection of a large number of reflected waves, and thus exhibits a multipath propagation phenomenon. This makes the received signal's phase random and hard to predict, and can possibly lead to large fluctuation in the received power. Assuming that the received signal on each path is random, the model developed in [18] is based on a continuous time description of the scattered electric field received at a stationary receiver with multipath reception along N paths, expressed as $\epsilon^{(N)}(t) = \sum_{k=1}^{N} a_k \exp[i\varphi_k(t)]$, where a_k is the amplitude of the received signal along path k and i is the square root of -1. The phase factors $\exp[i\varphi_k(t)]$ are independent and uniformly distributed on a unit circle in the complex plane and for each t. In addition, it is assumed that the phase $\varphi_k(t)$ satisfies the following SDE: $d\varphi_k(t) = B^{\frac{1}{2}} dW_k(t)$, where $\varphi_k(0)$ are uniformly distributed on $[0, 2\pi)$, W_k are independent Weiner processes, and B is a constant that represents the rate of change in the phase of the received signal. By integrating the above SDE it is readily seen that $\mathrm{VAR}(\varphi_k(t) - \varphi_k(0)) = Bt$.

Using stochastic calculus, it was established in [18] that the amplitude process is given by $\Psi(t) = I(t) + iQ(t)$, where $I(t)$ and $Q(t)$ are the in phase and quadrature components of the incoming waves received at time t, and can be represented by the following two SDEs:

$$dI(t) = -\frac{1}{2} BI(t)dt + \frac{\sqrt{2}}{2}\sigma B^{\frac{1}{2}}\, dW^{(I)}(t); \tag{1}$$

$$dQ(t) = -\frac{1}{2} BQ(t)dt + \frac{\sqrt{2}}{2}\sigma B^{\frac{1}{2}}\, dW^{(Q)}(t), \tag{2}$$

with $I(0) = 0$, $Q(0) = 0$, and $W^{(I)}(t)$ and $W^{(Q)}(t)$ two independent standard Wiener processes. The parameter B makes these two SDEs *mean-reverting*, i.e., the process, in equilibrium, approaches the mean [19]. Such processes are also referred to as Ornstein-Uhlenbeck processes [19]. The parameter σ^2 represents the stationary magnitude of the scattering power averaged over an asymptotically large number of propagation paths [18], and is shown to be the asymptotic (in t) variance of ϵ_t and satisfies: $\sigma^2 = \sum_{k=1}^{\infty} a_k^2$, which is assumed to be finite, assuming no single path dominates: $\lim_{N\to\infty} \frac{a_j^2}{\sum_{k=1}^{N} a_k^2} = 0$.

The above summarizes what was developed in [18]. We now proceed to derive the power process received at the receiving antenna at time t. This is given by $R(t) = \sqrt{I^2(t) + Q^2(t)}$. Assuming processes (1) and (2), and using standard arguments from stochastic calculus [19], we have the following lemma:

Lemma 1. *Let $\bar{I}(t) = e^{\frac{1}{2}Bt} I(t)$ and $\bar{Q}(t) = e^{\frac{1}{2}Bt} Q(t)$. Then*

$$d\bar{I}(t) = \frac{\sqrt{2}}{2}\sigma B^{\frac{1}{2}} e^{\frac{1}{2}Bt} dW^{(I)}(t), \tag{3}$$

$$d\bar{Q}(t) = \frac{\sqrt{2}}{2}\sigma B^{\frac{1}{2}} e^{\frac{1}{2}Bt} dW^{(Q)}(t). \tag{4}$$

Furthermore, both $\bar{I}(t)$ and $\bar{Q}(t)$ are normally distributed, with mean 0 and variance $\frac{\sigma^2}{2}(e^{Bt} - 1)$.

Proof. Let $f(x,t) = e^{\frac{1}{2}Bt}x$. Then $f_t = \frac{\partial f}{\partial t} = \frac{1}{2} B e^{\frac{1}{2}Bt} x$, $f_x = \frac{\partial f}{\partial x} = e^{\frac{1}{2}Bt}$ and $f_{xx} = \frac{\partial^2 f}{\partial^2 x} = 0$. Using Ito's formula (Theorem 4.2.3 [19]) and replacing x with $I(t)$, we have

$$\begin{aligned} d\bar{I}(t) &= f_t dt + f_x dI(t) + \frac{1}{2} f_{xx} (dI(t))^2 \\ &= \frac{1}{2} B e^{\frac{1}{2}Bt} I(t)dt + e^{\frac{1}{2}Bt}\left(-\frac{1}{2} BI(t)dt \right. \\ &\quad \left. + \frac{\sqrt{2}}{2}\sigma B^{\frac{1}{2}}\, dW^{(I)}(t)\right) \\ &= \frac{\sqrt{2}}{2}\sigma B^{\frac{1}{2}} e^{\frac{1}{2}Bt} dW^{(I)}(t), \end{aligned}$$

where we have substituted Eq. (1) for $dI(t)$. The second, Eq. (4), can be obtained by the same argument. Integrating, say the equation for $\bar{I}(t)$, we see that

$$\bar{I}(t) = \bar{I}(0) + \frac{\sqrt{2}}{2} B^{\frac{1}{2}}\sigma \int_0^t e^{\frac{1}{2}Bs} dW^{(I)}(s).$$

Using the facts that the mean of the Ito's integral $\int_0^t e^{\frac{1}{2}Bs} dW^{(I)}(s)$ is zero and its variance $\int_0^t e^{Bs} ds$ we get the result. □

An immediate consequence of the above lemma is that

$$\bar{R}(t) = \sqrt{\bar{I}(t)^2 + \bar{Q}(t)^2} \tag{5}$$

has a Rayleigh distribution with parameter $\sqrt{\frac{\sigma^2}{2}(e^{Bt} - 1)}$. The main theorem is given as follows:

Theorem 1. *The power process $R(t)$ satisfies the following SDE:*

$$dR(t) = -\frac{BR(t)}{2} dt + \frac{1}{4}\frac{B\sigma^2}{R(t)} dt + \frac{1}{\sqrt{2}}\sigma B^{\frac{1}{2}}\, dW(t) \tag{6}$$

with $R(0) = r_0$, some constant, and W a standard Wiener process.

Proof. Consider $\bar{R}(t)$ as in Eq. (5) and note that $\bar{R}(t) = e^{\frac{1}{2}Bt} R(t)$. Now, consider the function $f(x,y) = \sqrt{x^2 + y^2}$, for which we have the following first and second order partial derivatives: $f_x = \frac{x}{\sqrt{x^2+y^2}}$, $f_y = \frac{y}{\sqrt{x^2+y^2}}$, $f_{xx} = \frac{1}{\sqrt{x^2+y^2}} - \frac{x^2}{(\sqrt{x^2+y^2})^3}$, $f_{yy} = \frac{1}{\sqrt{x^2+y^2}} - \frac{y^2}{(\sqrt{x^2+y^2})^3}$ and $f_{xy} = -\frac{xy}{(\sqrt{x^2+y^2})^3}$.

Substituting the above into Ito's formula when differentiating (5), and using standard results on Wiener processes: $dW^{(I)}(t)^2 = dt, dW^{(Q)}(t)^2 = dt, dW^{(I)}(t)dW^{(Q)}(t) = 0$, we get

$$d\bar{R}(t) = \frac{\bar{I}(t)d\bar{I}(t)}{\bar{R}(t)} + \frac{\bar{Q}(t)d\bar{Q}(t)}{\bar{R}(t)} + \frac{1}{4}\sigma^2 B e^{Bt}\frac{1}{\bar{R}(t)}dt. \quad (7)$$

Consider the first two terms in the above expression. It is seen

$$\frac{\bar{I}(t)d\bar{I}(t)}{\bar{R}(t)} + \frac{\bar{Q}(t)d\bar{Q}(t)}{\bar{R}(t)} = \frac{\sqrt{2}}{2}\sigma B^{\frac{1}{2}} e^{\frac{1}{2}Bt}\left[\frac{I(t)dW^{(I)}(t) + Q(t)dW^{(Q)}(t)}{R(t)}\right].$$

We have that

$$\frac{I(t)dW^{(I)}(t) + Q(t)dW^{(Q)}(t)}{R(t)} = \frac{1}{R(t)}[I(t), Q(t)]\begin{bmatrix} dW^{(I)}(t) \\ dW^{(Q)}(t)\end{bmatrix}$$

and by the definition of $R(t)$

$$\frac{1}{R(t)^2}[I(t), Q(t)]\begin{bmatrix} I(t) \\ Q(t)\end{bmatrix} = 1.$$

Therefore, using Theorem 8.4.2 of [19] we conclude that $\int_0^t \frac{I(s)dW^{(I)}(s) + Q(s)dW^{(Q)}(s)}{R(s)}ds$ has the same law as a Wiener process, denoted as $W(t)$, independent of $W^{(I)}$ and $W^{(Q)}$. This means that we can write

$$\frac{\bar{I}(t)d\bar{I}(t)}{\bar{R}(t)} + \frac{\bar{Q}(t)d\bar{Q}(t)}{\bar{R}(t)} = \frac{\sqrt{2}}{2}\sigma B^{\frac{1}{2}} e^{\frac{1}{2}Bt}\, dW(t).$$

Substituting the above into (7) we obtain:

$$d\bar{R}(t) = \frac{\sqrt{2}}{2}\sigma B^{\frac{1}{2}} e^{\frac{1}{2}Bt} dW(t) + \frac{1}{4}\sigma^2 B e^{Bt}\frac{1}{\bar{R}(t)}dt. \quad (8)$$

Since $R(t) = e^{-\frac{1}{2}Bt}\bar{R}(t)$, we have

$$dR(t) = -\frac{1}{2}Be^{-\frac{1}{2}Bt}\bar{R}(t)dt + e^{-\frac{1}{2}Bt}d\bar{R}(t). \quad (9)$$

Replacing $d\bar{R}(t)$ with (8) in the above equation gives the desired result. □

The power process we propose to use in this paper is (6) with a modified mean reverting term:

$$dR(t) = \frac{B}{2}(\mu - R(t))dt + \frac{1}{4}\frac{B\sigma^2}{R(t)}dt + \frac{1}{\sqrt{2}}\sigma B^{\frac{1}{2}} dW(t). \quad (10)$$

The additional term $\frac{B\mu}{2}dt$ on the RHS, which is now part of the mean reverting term, has the effect of steering the mean of the process to approximately μ (at $t \to \infty$). To summarize, our model given in (10) consists of three terms: first a "mean reverting" process (or the O-U process) with mean μ, the second the "radial" term, and the third a volatility term (together a Bessel process). Using the terminology of [20], we will call this process a *Radial Ornstein-Uhlenbeck process*.

Three unknown parameters uniquely define this model: μ, B, and σ. μ will be referred to as the mean. B will be referred to as the phase constant; it corresponds to the rate at which the received signal phase changes. σ^2 will be referred to as the power constant (not to be confused with the received power); it is the sum of signal magnitudes over all paths. The value $B\sigma^2$ determines the volatility of this process. In the next section, we will show how these three parameters can be estimated using spectrum measurement data for training.

2.2 Parameter Estimation

In order to estimate the three unknowns μ, B and σ from real measurement data, we first rearrange terms in (10) to obtain the following:

$$\frac{dW(t)}{\sqrt{dt}} = \frac{\sqrt{2}}{\sigma B^{\frac{1}{2}}}\left\{\frac{dR(t)}{\sqrt{dt}} - \frac{B}{2}(\mu - R(t))\sqrt{dt} - \frac{B\sigma^2}{4R(t)}\sqrt{dt}\right\}. \quad (11)$$

Note that the left hand side of the above equation is now a zero-mean, unit-variance normally distributed random variable. The idea behind our parameter estimation procedure is to use real measurement data to generate data points corresponding to the right hand side of Eq. (11), and then match the first two (or more) sample moments to that of the 0-mean unit-variance normal distribution, thereby solving three unknowns (μ, B, σ).

Specifically, for a given frequency band our measurements are in the form of a time series of power readings, denoted as $\hat{R}(t_i)$, $i = 0, 1, 2, \ldots, N$. From these measurements we can now obtain successive differences between these readings, denoted as $d\hat{R}(t_i) = \hat{R}(t_i) - \hat{R}(t_{i-1})$, $i = 1, 2, \ldots, N$. We can also obtain the differences in sampling times, denoted as $dt_i = t_i - t_{i-1}$, $i = 1, 2, \ldots, N$. For our measurement data, sampling times are evenly spaced. Therefore in our case dt_i is treated as a constant.

Following this, the original measurement data may be viewed as a collection of N triples $(\hat{R}(t_i), d\hat{R}(t_i), dt_i)$, $i = 1, 2, \ldots, N$. Each such triple will now be referred to as a *sample* within the context of estimation and testing. From this collection of samples, we now select a random subset $\mathcal{N}_{est}$ of size N_{est} for estimation. We plug in each selected sample into the RHS of Eq. (11) and obtain the following data point $\hat{w}_i$ for $i \in \mathcal{N}_{est}$:

$$\hat{w}_i = \frac{\sqrt{2}}{\sigma B^{\frac{1}{2}}}\left\{\frac{d\hat{R}(t_i)}{\sqrt{dt_i}} - \frac{B}{2}(\mu - \hat{R}(t_i))\sqrt{dt_i} - \frac{B\sigma^2}{4\hat{R}(t_i)}\sqrt{dt_i}\right\}. \quad (12)$$

This gives us a total of N_{est} data points $\{\hat{w}_i, i \in \mathcal{N}_{est}\}$, each a function of μ, B and σ. We also obtain an estimate of the mean by $\hat{\mu} = \frac{1}{N_{est}}\sum_{i\in\mathcal{N}_{est}}\hat{R}(t_i)$. These three unknown parameters can now be estimated by matching (1) the sample mean of the data set $\{\hat{w}_i, i \in \mathcal{N}_{est}\}$ to 0; (2) its sample variance to 1; (3) the parameter μ to the mean estimate $\hat{\mu}$; That is, the parameters are estimated by matching the first two sample moments to the first two moments of the 0-mean unit-variance normal distribution and matching the parameter μ to the estimated mean of the received power.

In our experiments, we obtain the estimates by solving the following minimization problem: $\min_{\mu,B,\sigma}(m_1 - 0)^2 + (m_2^{1/2} - 1)^2 + \frac{(\mu-\hat{\mu})^2}{\mu^2} + P$ where m_i denotes the ith sample moment of the data set $\{\hat{w}_i, i \in \mathcal{N}_{est}\}$, and 0 and 1 are the first two moments of the standard Normal distribution. P is

a penalty term designed to penalize the minimization when the parameter B is negative or too close to zero (note that in the process $R(t)$, B is a positive term). The term is in the form of $P = \mathcal{C}(B-\delta)^2$ when $B < \delta$ and $P = 0$ when $B \geq \delta$ (where δ is a small number and $\delta, \mathcal{C} > 0$).

Once these parameters are estimated, we use the remaining $N - N_{est}$ samples for testing and model verification. This is done in a very similar way as in estimation. Specifically, the testing samples are also plugged into the RHS of Eq. (11). However, this time the computation is done with the estimated values of μ, B and σ. This gives us $N - N_{est}$ data points, also commonly referred to as the *residual of the test data*. Our first model verification test in Section 3 lies in checking whether the residual follows the standard normal distribution.

2.3 Analytical Expression of SDE Distribution

An immediate application of the SDE model is to derive received power distribution in a channel. As mentioned earlier, the SDE model belongs to the family of diffusion models. Diffusion models are often used in large-scale queuing systems as good alternatives to the discrete valued Markov chains. Specifically, by allowing the queue to have continuous values, the discrete valued Markov chain can be approximated by a diffusion model with the appropriate parameters. This makes it feasible to derive the steady state distribution of the queue analytically, which is otherwise impossible under a discrete model. Below we use similar techniques to obtain the steady state distribution of the power process.

For any stochastic differential equation of the form

$$dX_t = U(X_t, t)dt + \sqrt{2D(X_t, t)}dW_t, \tag{13}$$

the distribution of the process $f(x,t)$, where $f(x,t)$ denotes the probability $P(X_t = x)$, satisfies the following equality by using the Fokker-Planck equation [14]:

$$\frac{\partial}{\partial t} f(x,t) = -\frac{\partial}{\partial x}[U(x,t)f(x,t)] + \frac{\partial^2}{\partial x^2}[D(x,t)f(x,t)]. \tag{14}$$

In our SDE model, $R(t)$, the power process, takes the role of X_t in Eq. (13). Recalling for convenience the model:

$$dR(t) = \frac{B}{2}(\mu - R(t)))dt + \frac{B\sigma^2}{4R(t)}dt + \frac{\sigma\sqrt{B}}{\sqrt{2}}dW(t), \tag{15}$$

we then have the following mapping between (13) and (15): $U(X_t,t) = \frac{B}{2}(\mu - X_t) + \frac{B\sigma^2}{4x}$, $D(X_t,t) = \sigma^2 B/4$. Assuming that the process reaches a steady state, i.e., $\frac{\partial}{\partial t} f(x,t) = 0$. Then integrating the right hand side and suppressing the argument t will give us

$$\frac{B}{2}\left(\mu - x + \frac{\sigma^2}{2x}\right) f(x) = \frac{\sigma^2 B}{4} f'(x), \tag{16}$$

or

$$\frac{2}{\sigma^2}(\mu - x) + \frac{1}{x} = \frac{f'(x)}{f(x)}. \tag{17}$$

Integrating the above over x we get

$$\frac{2}{\sigma^2}(\mu x - x^2/2) + \log x = \log f(x) + c. \tag{18}$$

Rearranging to solve for $f(x)$ we get,

$$f(x) = c \cdot x \cdot \exp\left(\frac{2}{\sigma^2}(\mu x - x^2/2)\right), \tag{19}$$

and the normalizing constant value $c = 1/\int_0^\infty x \cdot \exp\left(\frac{2}{\sigma^2}(\mu x - x^2/2)\right)$. This gives us the complete description of the steady state distribution function of the power process $R(t)$.

2.4 Synthesizing Spectrum Data

An important reason for developing any channel model is to provide a way to generate synthetic channel data (sample paths of energy levels in a channel) that are statistically close to variations observed in a real channel, so that one can easily generate a realistic "spectrum environment" in which to test and evaluate various algorithms and protocols. Below we show such synthetic data can be easily generated under our SDE model.

Taking Eq. (10) and integrating over an interval ϵ,

$$\begin{aligned} R(t+\epsilon) - R(t) &= \frac{B\mu}{2}(t+\epsilon-t) - \frac{B}{2}\int_t^{t+\epsilon} R(\tau)d\tau \\ &\quad + \frac{1}{4}B\sigma^2\int_t^{t+\epsilon}\frac{1}{R(\tau)}d\tau \\ &\quad + \sigma\frac{1}{\sqrt{2}}B^{\frac{1}{2}}(W(t+\epsilon) - W(t)) \\ &\approx \frac{B\mu}{2}\epsilon - \frac{B}{2}R(t)\epsilon + \frac{1}{4}B\sigma^2\frac{1}{R(t)}\epsilon \\ &\quad + \sigma\frac{1}{\sqrt{2}}B^{\frac{1}{2}}(W(t+\epsilon) - W(t)), \end{aligned} \tag{20}$$

where the approximation holds when ϵ is sufficiently small. Assuming we start from some initial condition $R(t_o)$ at time t_o, we can generate a sequence of values $R(t_o + k\epsilon)$ at times $t_o + k\epsilon$ for $k = 1, 2, \ldots$ with time resolution (or time step) of ϵ as follows. Note that $W(t+\epsilon) - W(t)$ is normally distributed with zero mean and variance ϵ.

1) Generate a random sample from the 0-mean ϵ-variance normal distribution; denote it by $W(t_o+\epsilon) - W(t_o)$.
2) Take this sample value into Eq. (20), replacing the corresponding part in the last term on the RHS.
3) Compute the RHS, which gives the difference between $R(t_o+\epsilon)$ and $R(t_o)$, hence we have generated a value for $R(t_o+\epsilon)$.
4) Repeat the above procedure indefinitely to produce a time series of desired length.

The end result of this procedure is a sequence of synthesized $R(t)$ values, representing a particular realization.

3 Model Verification Using Spectrum Measurement Data

In this section we employ a variety of techniques to verify the accuracy of the SDE model in capturing features of the

TABLE 1
Data Sets for Model Verification

Data Set	MHz	Primary user	Source
1	551	TV	CRAWDAD [15]
2	629	TV	CRAWDAD [15]
3	665	TV	CRAWDAD [15]
4	518	TV	measurement study [5]
5	738	TV	measurement study [5]
6	1,842	GSM	measurement study [5]

real data. We use spectrum data from two sources: (1) The first is our measurement study reported in [5], which was done over a period of multiple days continuously, and simultaneously at multiple locations. The resolution of the measurement is such that one energy reading (in μV) is produced for each band of width 200 KHz, for roughly every 75 seconds of sweep time over the range of 20 MHz to 3 GHz. (2) The second is a data set obtained from crawdad.org [15]. Compared to the first data set, this set consists of readings for a larger band (10 MHz), but over a much smaller sweep time of roughly 80 nano seconds. These two sets of data thus represent two very different measurement regimes: the first has low time resolution (sampling rate) but high spectral resolution (narrow bandwidth) while the second has a much higher sampling rate but wider bandwidth.

From these two data sets we selected the bands of TV Digital and TV Analog for training and verifying the SDE model. The center frequencies of these bands are listed in Table 1.

3.1 Q-Q Plots

We first directly check whether the residual follows a standard normal distribution, and in particular to check how far it is from the standard normal distribution, we employ the Quantile-Quantile(Q-Q) plot, a commonly used graphical statistical tool, see e.g, [21], [22], [23]. The quantiles are points taken at regular intervals from the cumulative distribution function (CDF) of a random variable. The p-quantile for a random variable X is the value x such that $\text{Prob}(X < x) = p$. A Q-Q plot shows the quantiles of a first data set against the quantiles of a second data set, and is therefore an intuitive (as well as visual) and efficient way to determine if two data sets follow a common distribution.

For two random data sets S_1 and S_2, the Q-Q plot is generated by first sorting each set in increasing order, and then sequentially placing points on the plot. The ith point is placed at coordinates (s_i^1, s_i^2), where s_i^1 and s_i^2 are the values of the ith data point in the sorted sets S_1 and S_2, respectively.

In order to check whether a data set follows a standard normal distribution, we will make the first data set the theoretical quantiles of the standard normal distribution and the second data set the residual of the test data on the Q-Q plot. If the points fall along a 45-degree reference line, then it is strong evidence that the residual follows the standard normal distribution. If the points fall along a line, but not the reference line, then this suggests that the residual follows normal distribution but is not exactly standard.

Fig. 1 shows the above normality test results for the three frequencies tested; for each frequency (the corresponding

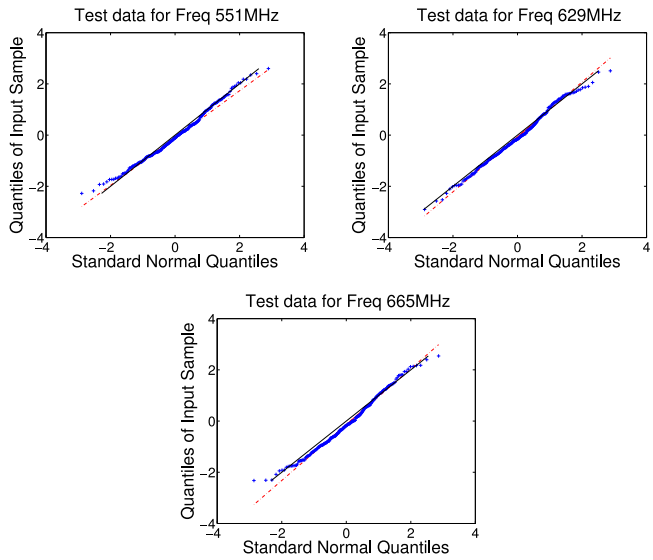

Fig. 1. Q-Q plot, 551 MHz (top left), 629 MHz (top right), 665 MHz (bottom).

data is now a collection of samples (triples) as described in the previous section), we randomly select 50 percent to be the estimation/training data set and the remaining 50 percent the testing data set.[2] The dashed line represents the best linear fit of the points (the residual)—the closer the points are to the dashed line, the more normally distributed the residual is. The solid line represents the 45-degree reference line. It can be seen that in all three cases the points have very good linear fit, indicating strong normality of the residuals. In addition, all three dashed lines are very close to the reference line, indicating the residuals follow close-to-standard normal distributions.

In Fig. 2 we show results on the next three frequencies.[3] As noted earlier, this data set has a much coarser time resolution, with 72 seconds between two consecutive data points (as opposed to 78 nano seconds of the first data set).

We observe that the data acquired from CRAWDAD and in [24] both fit really well but the CRAWDAD data fits better because the residuals in Fig. 1 fall more linearly on the 45-degree reference line. The reason is likely due to the fact that the first data set has a better time resolution. For the second data set, the fit is also good with the exception of 1,842 MHz.[4] Overall, these results verify that the SDE model can be applied to both sparse and fine sampling rates of data. Along with the fitting result, Table 2 shows the values of the estimated parameters; these values can be used in synthesizing spectrum utilization data, described in Section 2.4.

In the next few sections we further verify the SDE model by comparing the synthetic data with measured data, as additional, indirect means of validating the SDE model. It is

2. Randomly selecting a set for estimation is a standard procedure in statistical analysis; the resulting estimation represents the data better in case the underlying process is non-stationary.

3. These are reproduced from our previous study [24].

4. The reason that the model fails in this case is because when calibrating the data, we assumed that the whole set (Monday-Friday) comes from the same process which has the same three parameter values. Unfortunately, this set contains an outlier—one of the days has a mean significantly lower than the other four (see [24] for more detail).

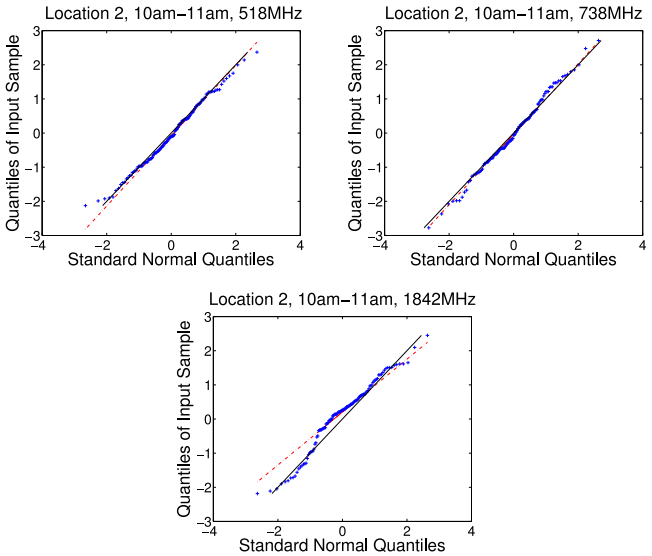

Fig. 2. Q-Q plot, location 2, 10-11 am: 518 MHz (top), 738 MHz (middle), and 1,842 MHz (bottom) from [24].

in general hard to directly compare sample paths simply by plotting them out. Instead, we first compare the entropy of sample paths generated through synthesis and from the measurement data, and then compare 3 time domain metrics: auto-correlation, level crossing rate, and the average fading duration (AFD) commonly used in fading analysis.

3.2 Entropy Comparison

This experiment is done by choosing a sliding window of size 10,000 (samples) from each of the sample paths. Within this window, we compute the power spectral density of the samples, from which we then compute the entropy of the same samples. We do this over the entire sequence with overlapping of 50 percent between windows.

Fig. 3 shows the results of the above comparison with measured data. It can be seen that the entropy measures of the two are very close in all cases (they also both remain steady throughout the entire sequence).

3.3 Autocorrelation and Power Spectral Density

We next take a window of size 1,000 (samples) and calculate the autocorrelation and power spectral density. We do the same for both the measurement data and the fitted SDE model. The results are shown in Figs. 4 and 5; for brevity we only show results for 551 and 629 MHz as the other results are similar. We see that the SDE-generated sample paths have a slightly higher autocorrelation in both frequencies, but the difference is quite minor and their shapes are

TABLE 2
Trained Parameters

Data Set	MHz	$\bar{\mu}$	b	σ
1	551	144.5060	1.2606e+07	93.1635
2	629	234.4078	1.5509e+07	124.3106
3	665	226.9313	1.7845e+07	145.8981
4	518	0.3785	0.0391	0.0327
5	738	0.9661	0.0383	0.0732
6	1,842	3488.0	0.0297	894.29

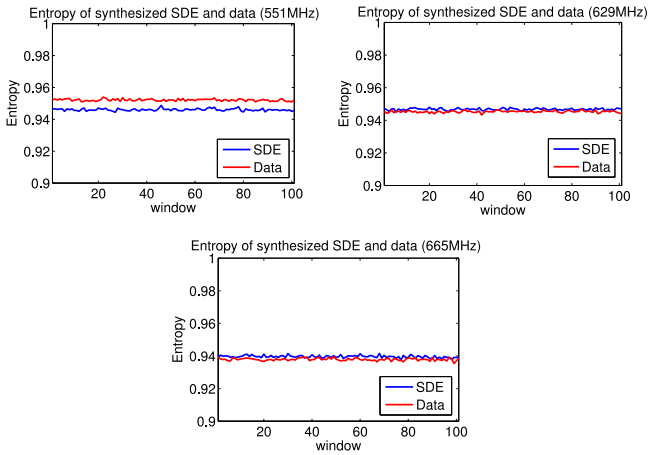

Fig. 3. Entropy of synthesized SDE and data (551 MHz top left, 629 MHz top right, 665 MHz bottom).

very similar. The power spectrum densities of both frequencies again appear to be very similar as shown in Fig. 5.

3.4 Level Crossing and Average Fading Duration

The level crossing rate is obtained by calculating how often the received power exceeds a certain power level, which is a measure of the rapidity of fading. We plot the level crossing rate of both the measured data and the synthetic data using SDE model fitted from the measured data. As shown in Fig. 6, the sample path for 551 MHz shows a very good fit where both the occurrence of the peak and the magnitude of the peak match very well. For 629 MHz, the position of the peak matches but the SDE model has a higher crossing rate than the actual measurement data.

The average fading duration calculates how long the received signal remains below a threshold. This indicates how often the receiver receives low signal from the transmitter. Fig. 7 shows the AFD of frequencies 551 and 629 MHz. Again

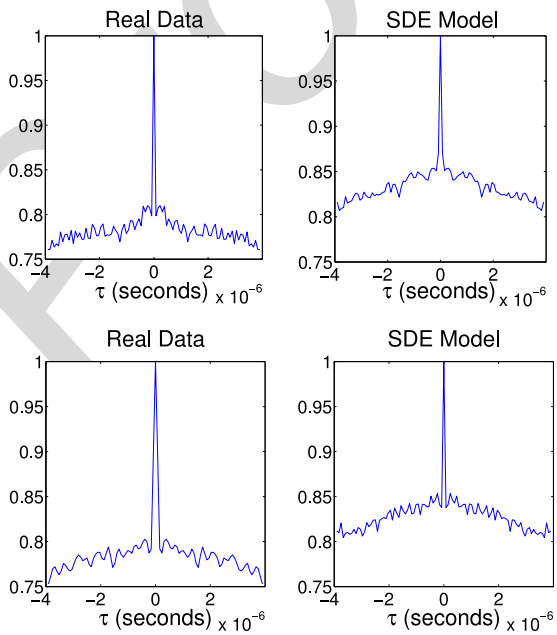

Fig. 4. Autocorrelation of synthesized SDE and measurement data (551 MHz top, 629 MHz bottom).

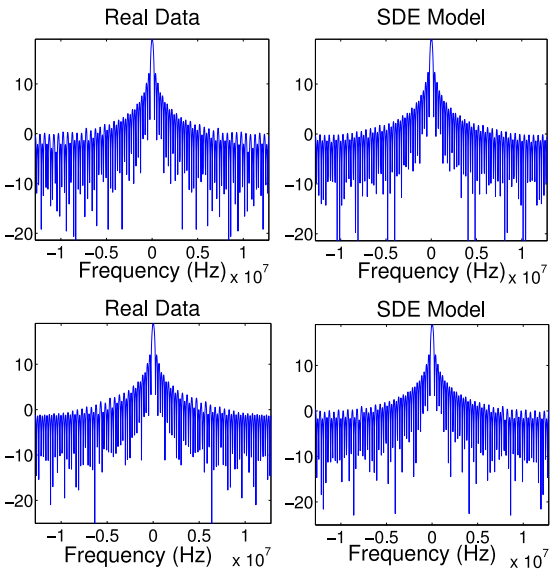

Fig. 5. Power spectral density of synthesized SDE and measurement data (551 MHz top, 629 MHz bottom).

we see the AFD of 551 MHz shows a better fit for the SDE model, though both are quite close to the actual AFD.

4 SDE As a Generalization of the GE Model

In this section we show that the SDE model can be used to generate a two-state GE model; furthermore, this GE model is virtually identical to that generated directly from the data. This in some sense validates the synthetic data generated by the SDE model through yet another means, and at the same time shows that the SDE model may be viewed as a more general modeling framework.

4.1 Generating the GE Model

Given a sample path (measured or synthesized) of continuous power levels, the discrete GE model may be generated by quantizing/discretizing the data into binary values (where 0 represents a bad channel condition/state and 1 a good channel condition/state). Using guidelines from [25], this discretization is done by viewing the channel as occupied when the received power exceeds the minimum received power (observed in the same sample path) by more than 3 dB. Both the measurement data and the synthetic data are already discrete in time, but can be easily down-sampled to obtain a desired time resolution.

We then compute the ratio between how many times BB and BG (respectively GG and GB) transitions occurred over the quantized sample paths, and use this as the estimate for

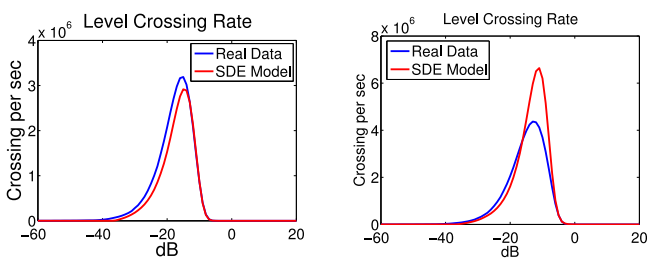

Fig. 6. Level crossing rate of synthesized SDE and measurement data (551 MHz left, 629 MHz right).

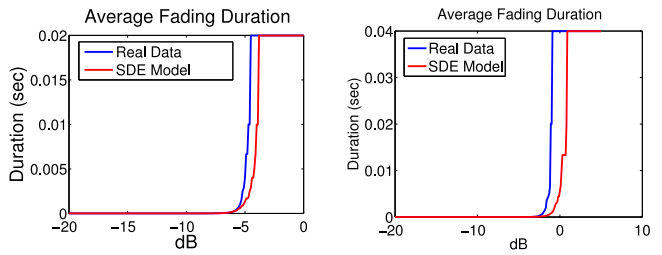

Fig. 7. Average fading duration of synthesized SDE and measurement data (551 MHz left, 629 MHz right).

the transition probability ratio P_{BB}/P_{BG} (respectively P_{GG}/P_{GB}). This ratio together with the fact the $P_{BG} = 1 - P_{BB}$ (respectively $P_{GB} = 1 - P_{GG}$) lead to the values of the transition probabilities for the two-state GE model, illustrated in Fig. 8.

4.2 Comparison between Data-Generated GE and SDE-Generated GE

The following experiment is performed on the sample paths collected at 551, 629, and 665 MHz, respectively:

1) Train the parameters of the SDE using this data set.
2) Synthesize data from this trained SDE model.
3) Quantize the synthesized data trace to 0-1 binary value as described above.
4) Obtain a GE model from the discretized synthetic data also as described above.
5) Obtain a second GE model directly from the original set of sample paths, similarly quantized.

We then compare these two GE models, the idea being that if the two are similar, then the SDE model can also be used to generate a valid GE model. The comparison results are listed in Table 3.

We see that the two GE models share very similar parameters. This indicates two things. First, since the two GE models are generated from the synthetic and the measurement data, respectively, this similarity suggests that the synthetic and the measurement data are very similar in nature, thereby validating the SDE model as a means of generating realistic synthetic data. Second, this means that we can use the SDE model to generate essentially equivalent GE models without having to rely on the original measurement data, i.e., once the SDE model has been trained, we only need to record the three parameters μ, b, σ for future use. In this sense the SDE model may be viewed as a more general modeling framework.

It is also worth noting that the GE model, i.e., its parameters, can be calculated directly from the SDE model, given the threshold, Th, used to quantize the data. Specifically, since we know the steady state distribution from Section 2.3, by observing that the SDE model consists of a deterministic term and a normal distribution,

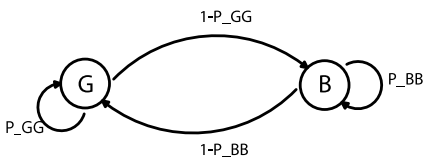

Fig. 8. The Gilbert-Elliott (two-state) model.

TABLE 3
Data Sets for Model Verification

Data Set	MHz	GE from data		GE from SDE	
		P_{BG}	P_{GG}	P_{BG}	P_{GG}
1	551	0.6489	0.6494	0.6350	0.6201
2	629	0.6506	0.6499	0.6427	0.6221
3	665	0.6151	0.6439	0.6615	0.5970

we can obtain the following expression for the transition probability:

$$P_{GG} = \int_0^{Th} f(x)\Phi\left(\frac{\sqrt{2}}{\sigma\sqrt{B}}\left(\frac{Th-x}{dt}+\frac{B}{2}(x-\mu)-\frac{B\sigma^2}{4x}\right)\right)dx, \tag{21}$$

i.e., P_{GG} is calculated by integrating the distribution of the power process being below the threshold multiplied by the probability that the next step falls below the threshold. The same can be done for P_{BB}. This method allows us to directly obtain the GE model from the SDE model without having to go through the synthetic data. This method can be extended to any N-state Markov model and will be used in Section 5.

4.3 Fitting Performance of the GE Model

Following the previous section, it would be natural to question how well the GE model fits the data if we only consider the quantized, binary description of the channel. In this section we examine this by comparing (1) the autocorrelation of the sample paths generated by the GE model and the quantized actual data, and (2) the run length distribution of error and error-free runs/sequences observed in these samples.

We first compare the autocorrelation and power spectrum density of the actual data (data set 1, quantized) and the sample paths generated by the GE model. These are shown in Figs. 9 and 10, respectively.

We see that the autocorrelation and power spectral density of the GE model match the actual discretized data samples very well with similar maximum values and similar shapes.

Next we examine the length of consecutive available states (consecutive '1's) and consecutive unavailable states (consecutive '0's). Fig. 11 (left) shows the distribution of consecutive runs of availability. The x-axis is the run length plotted in logarithm scale, and the y-axis is the proportion of runs of that length. We observe that the GE model

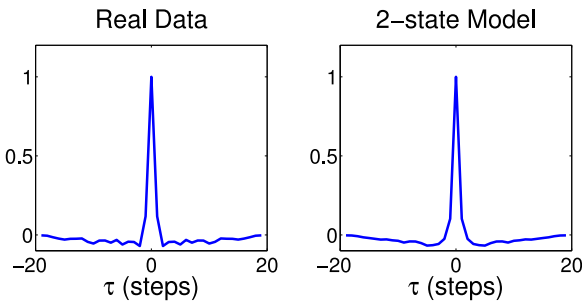

Fig. 9. The autocorrelation of real data and the GE model over a 300-second duration.

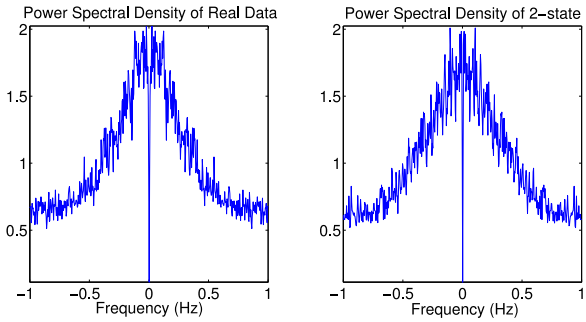

Fig. 10. The power spectral density of real data and the GE model over a 300-second duration.

generates run lengths of both error and error-free close to the actual data.

The autocorrelation and the error length show that the GE model fits the actual data really well, suggesting that the GE model is a good choice when the discretized binary representation of the channel condition is sufficient. In other words, if an application only needs to observe whether the channel is in a good quality, the GE model is more or less adequate. This however is no longer the case if we require higher resolution. In the next section, we demonstrate that when we need more information than just discretized binary values, the SDE model provides much more richness.

5 Predictive Performance of the SDE Model

One obvious advantage of having a continuous model like SDE is the richness of the data it provides compared to a discrete (esp. binary) model like GE. In this section we take a closer look at this aspect within the context of the respective model's ability to predict channel conditions. In many applications including channel-aware transmission scheduling [26] and opportunistic spectrum access [27], [28], collecting and predicting channel condition (or channel side information (CSI) in some contexts) is often critical to the effectiveness of the overall mechanism. For these applications, we often wish to obtain information more than just the binary value representing whether the channel is occupied. This is so that we could, for instance, more accurately calculate the SNR from the interference power to estimate the achievable transmission throughput.

Below we examine a channel model's ability to predict, in a discrete-time setting, the received power of the next step given the current power level. We do so for the SDE model, and an N-state Markov model, a more general version of

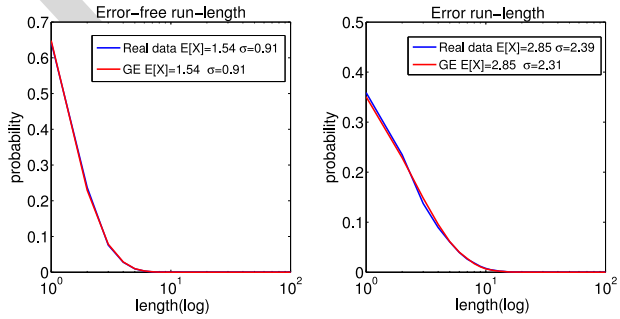

Fig. 11. Comparison between the SDE and GE models: consecutive available run length (left) and consecutive occupied run length (right).

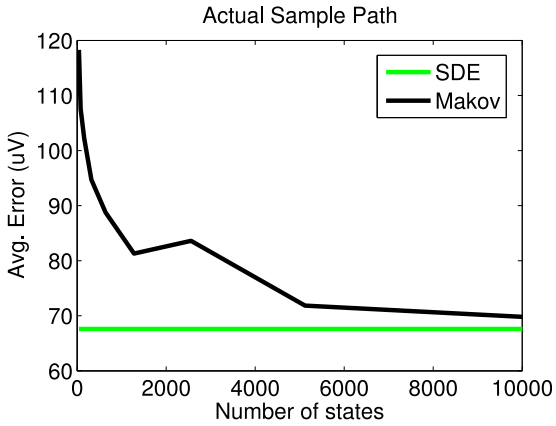

Fig. 12. Average error obtained using Markov models of different number of states compared with SDE.

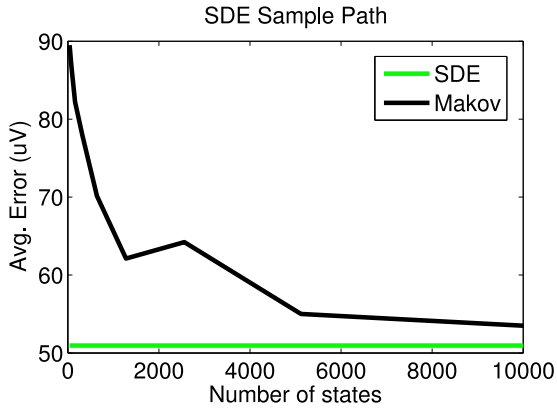

Fig. 13. Average error obtained using Markov models of different number of states compared with SDE.

the two-state GE model which allows us to adjust the resolution (the size of the state space N) and the corresponding quantization error.

We use the data-generated SDE model as a channel prediction model and compare it with the prediction of an N-state Markov model. Since both models have Markovian behavior, knowing the current state is sufficient to predict the next state. The error of the prediction model is defined as the absolute value of the difference between the prediction and the actual value. Given the past values of a process, the SDE model satisfies the Markovian property and predicts the next time step purely on the current power value:

$$R(t+\epsilon) \approx R(t) + \frac{B\mu}{2}\epsilon - \frac{B}{2}R(t)\epsilon + \frac{1}{4}B\sigma^2\frac{1}{R(t)}\epsilon + \sigma\frac{1}{\sqrt{2}}B^{\frac{1}{2}}(W(t+\epsilon) - W(t)). \quad (22)$$

Notice that only the last term contributes to a stochastic value with mean 0 and variance ϵ; all other terms are deterministic. If our metric is the absolute loss $|$prediction value—actual value$|$, then the best prediction is to predict the expected value. The expected value is obtained by discarding the stochastic terms and leaving only the deterministic terms. Thus, in this comparison we will use $R(t+\ \epsilon) = R(t) + \frac{B\mu}{2}\epsilon - \frac{B}{2}R(t)\epsilon + \frac{1}{4}B\sigma^2\frac{1}{R(t)}\epsilon$ as the SDE model prediction.

For the N-state Markov model, we take the first 200,000 sample points from the CRAWDAD data for training. The N-state Markov model requires a quantization of different power levels. The quantization level is determined by dividing the sorted 200,000 sample points into 200,000/N levels; all points are quantized into the mean value of the level it is placed in. We can then construct the N-state transition probability matrix from the transitions of the 200,000 data points between the quantile levels.

It is worth noting that the SDE model only requires several hundred of sample points for training to obtain the three parameters (μ, b, σ) while the N-state Markov model requires at least two orders of magnitude (100X) more in the number of training samples. Also, the N-state Markov model requires $N \times N$ storage to hold the transition probability matrix.

Fig. 12 shows the average prediction error of the N-state Markov model for $N \in [1, 10, 000]$. Compared with the SDE model, we can see that the N-state Markov model is not very useful in prediction when N is small, but can achieve predictions close to the SDE model by trading space for accuracy. Fig. 13 shows the same result but replacing CRAWDAD data by SDE generated sample path. As expected, since it is actually generated from the SDE model, the SDE model predicts the power levels even more accurately. This error value is actually the average deviation of the last term in equation (22).

We next show the prediction performance over multiple time steps. We fix the number of states for the Markov model, but vary the prediction steps n from 1 to 37. The prediction of the Markov model is done by multiplying the transition matrix of the Markov model by it self for n times. Assuming the trained transition matrix is correct, this will give the correct n step transition probability matrix. Prediction under the SDE is done by recursively plugging in Equation (22) for n times. The results are shown in Fig. 14.

We see that the SDE clearly outperforms under all step sizes. The error accumulates when we increase the prediction steps, but this saturates to a near-constant value when the step is larger than 5. This is most likely because after five steps the dependency of the future power level on the current value becomes negligible, so that the prediction of both models become the steady state value.

It should be mentioned that one of the main reasons that contributed to the error under the Markov prediction model is the training process it takes to obtain the transition matrix—the measurement data we have appears insufficient in volume for this purpose when N is large. We thus conduct the following, improved experiment. We first use the SDE model to directly calculate the transition probabilities as shown earlier, so as to minimize this training error. This would be equivalent to training the Markov model using very long data traces. We then compare the prediction performance of the SDE and this more accurate N-state model over a synthetic sample path; the results are shown in Fig. 15. We see that the error of the Markov model is reduced even with small state space N. Note that when N is sufficiently large, the Markov model can predict slightly better than the SDE does. This is because under the Markov model we are truly calculating the probability of multi-step transitions

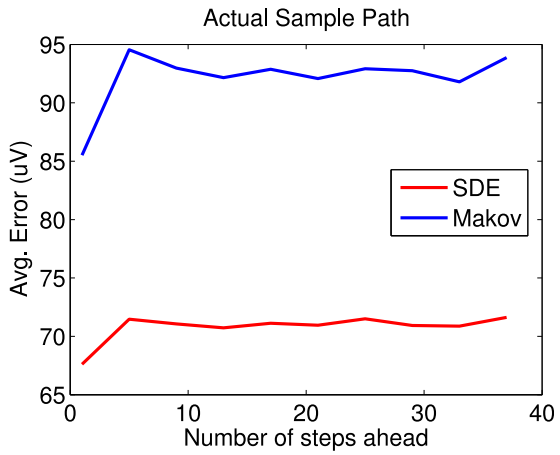

Fig. 14. Average error obtained using a 1,000-state Markov model and SDE for n-step prediction.

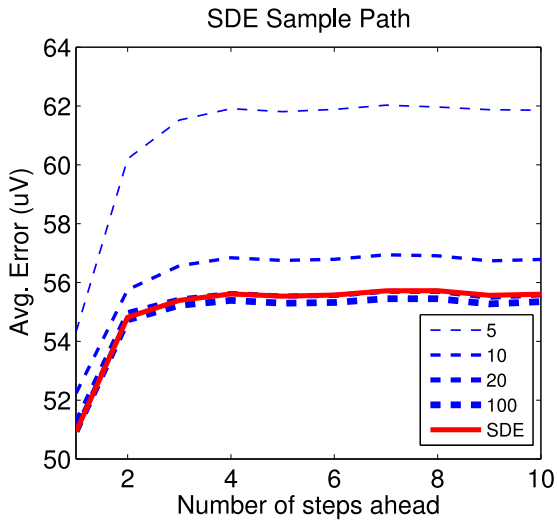

Fig. 15. Markov model obtained from SDE model.

whereas under the SDE model we recursively calculate the mean which is only an estimate. The improvement is, however, very minor, and comes at enormous computational expenses.[5]

It is worth highlighting a further advantage of the SDE model, which is flexibility. This can be seen in two aspects: (1) suppose the training data have measurements taken at nonuniform or random times. In this case it would be hard or even impossible to use this data to train a simple N-state Markov model because the transition of the Markov model is at a fixed time step. (2) If we want to change the discrete time unit of the Markov model, we would need to either retrain the model (in the case of up-sampling), or perform matrix multiplication (in the case of down-sampling and only feasible when the new time unit is an integer multiple of the current one). In contrast, the SDE model does not require a uniformly-sampled training data. Recall that in Section 2.2, the training is done with the triple $(\hat{R}(t_i), d\hat{R}(t_i), dt_i)$ where dt_i can be any real number. In addition, once we obtain the SDE model it is very easy to select/adjust the desired time resolution by choosing the Brownian motion term with the corresponding ϵ in Eq. (20) without having to retrain the model.

We end this section by commenting on how the SDE model and the N-state Markov model compare when used in simulation. Under the SDE model a sample path is generated using steps (1)-(4) outlined in Section 2.4. The complexity of these steps are quite minimal and comparable to that required when using the Markov model. One difference lies in the fact that we may need to down-sample the SDE sample path, as the choice of time step size ϵ may be smaller than a single discrete time step under the Markov model depending on the scenario being simulated. This can be achieved by "generate-and-discard": if the desired sample rate is once per $k\epsilon$ time units, then all intermediate values $R(t + i\epsilon)$ are discarded for $i = 1, 2, \ldots, (k-1)$. However, generating these unused data points should not be a significant burden computationally.

5. Again, note that we actually had to use the SDE model to generate the N-state model for lack of sufficient training, so this comparison would not have been possible without the SDE model!

6 Conclusion

In this paper we introduced a stochastic differential equation model to describe the secondary wireless channel power and compared it with N-state Markov models. We show that the SDE model fits spectrum measurement data very well under different measurement regimes. The SDE model can easily generate synthesized sample paths whose entropy measure is consistent with the original measurement data. The SDE model is a more general modeling framework that can be used to obtain an N-state Markov model. While we show that the two-state GE model is a good choice when binary representation of the channel condition is sufficient, the SDE model is in general much more accurate and easier to use than an N-state model.

Acknowledgments

This work is partially supported by the US NSF under grants CIF-0910765 and CNS-1217689, and the ARO under grant W911NF-11-1-0532. A part of the paper appeared in WiOpt'11. S.-P. Sheng is the corresponding author.

References

[1] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[2] M. A. McHenry, P. A. Tenhula, D. McCloskey, D. A. Roberson, and C. S. Hood, "Chicago spectrum occupancy measurements & analysis and a long-term studies proposal," in *Proc. 1st Int. Workshop Technol. Policy Accessing Spectrum*, 2006.

[3] M. A. McHenry, "NSF spectrum occupancy measurements project summary," Shared Spectrum Company Report, Aug. 2005.

[4] M. H. Islam, C. L. Koh, S. W. Oh, X. Qing, Y. Y. Lai, C. Wang, Y.-C. Liang, B. E. Toh, F. Chin, G. L. Tan, and W. Toh, "Spectrum survey in Singapore: Occupancy measurements and analyses," in *Proc. 3rd Int. Conf. Cognitive Radio Oriented Wireless Netw. Commun.*, May 2008, pp. 1–7.

[5] D. Chen, S. Ying, Q. Zhang, M. Liu, and S. Li, "Mining spectrum usage data: A large-scale spectrum measurement study," in *Proc. ACM Int. Conf. Mobile Comput. Netw.*, Sep. 2009, pp. 13–24.

[6] P. F. Marshall, "Closed-form analysis of spectrum characteristics for cognitive radio performance analysis," in *Proc. 3rd IEEE Symp. New Frontiers Dynam. Spectrum Access Netw.*, Oct. 2008, pp. 1–12.

[7] A. Kopke, A. Willig, and H. Karl, "Chaotic maps as parsimonious bit error models of wireless channels," in *Proc. 22nd Annu. Joint Conf. IEEE Comput. Commun.*, vol. 1, 2003, pp. 513–523.

[8] A. Konrad, B. Zhao, A. Joseph, and R. Ludwig, "A markov-based channel model algorithm for wireless networks," *Wireless Netw.*, vol. 9, no. 3, pp. 189–199, 2003.
[9] E. Gilbert, "Capacity of a burst-noise channel," *Bell Syst. Tech. J*, vol. 39, no. 9, pp. 1253–1265, 1960.
[10] E. Elliott, "Estimates of error rates for codes on burst-noise channels," *Bell Syst. Tech. J*, vol. 42, no. 9, pp. 1977–1997, 1963.
[11] J. McDougall and S. Miller, "Sensitivity of wireless network simulations to a two-state Markov model channel approximation," in *Proc. IEEE Global Telecommun. Conf.*, vol. 2, 2003, pp. 697–701.
[12] J. Hartwell and A. Fapojuwo, "Modeling and characterization of frame loss process in ieee 802.11 wireless local area networks," in *Proc. IEEE 60th Veh. Technol. Conf.*, vol. 6, 2004, pp. 4481–4485.
[13] Y. Yu and S. Miller, "A four-state Markov frame error model for the wireless physical layer," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2007, pp. 2053–2057.
[14] P. G. Harrison and N. M. Patel, *Performance Modelling of Communication Networks and Computer Architectures (International Computer S*. Reading, MA, USA: Addison-Wesley, 1992.
[15] J. Hoffbeck and A. Melton. (2008, Jul.). CRAWDAD data set up/rfrecordings (v. 2008-07-15). [Online]. Available: http://crawdad.cs.dartmouth.edu/up/rfrecordings
[16] T. R. Fields and R. J. A. Tough, "Stochastic dynamics of the scattering amplitude generating K-distributed noise," *J. Math. Phys.*, vol. 44, no. 11, pp. 5212–5223, Nov. 2003.
[17] C. D. Charalambous and N. Menemenlis, "A state-space approach in modeling multipath fading channels via stochastic differential equations," in *Proc. IEEE Int. Conf. Commun.*, vol. 7, Jun. 2001, pp. 2251–2255.
[18] T. Fang, T. R. Fields, and S. Haykin, "Stochastic differential equation theory applied to wireless channels," *IEEE Trans. Commun.*, vol. 55, no. 8, pp. 1478–1489, Aug. 2007.
[19] B. Oksendal, *Stochastic Differential Equations—An Introduction With Applications*. Berlin, Germany, Springer, 1998.
[20] A. N. Borodin and P. Salminen, *Handbook of Brownian Motion—Facts and Formulae*. Berlin, Germany, Birkhäuser, 2000.
[21] G. Blom, *Statistical Estimates and Transformed Beta Variables*. New York, NY, USA: Wiley, 1958.
[22] J. Chambers, W. Cleveland, B. Kleiner, and P. Tukey, *Graphical Methods for Data Analysis*. Belmont, CA, USA: Wadsworth, 1983.
[23] W. Cleveland, *The Elements of Graphing Data*. Hobart Press, 1994.
[24] L. Yang, S. Sheng, R. Saigal, M. Liu, D. Chen, and Q. Zhang, "A stochastic differential equation model for spectrum utilization," in *Proc. Int. Symp. Model. Optim. Mobile, Ad Hoc Wireless Netw.*, 2011, pp. 220–227.
[25] D. Chen, S. Yin, Q. Zhang, M. Liu, and S. Li, "Mining spectrum usage data: A large-scale spectrum measurement study," in *Proc. 15th Annu. Int. Conf. Mobile Comput. Netw.*, 2009, pp. 13–24.
[26] S. Ahmad, M. Liu, T. Javidi, Q. Zhao, and B. Krishnamachari, "Optimality of myopic sensing in multichannel opportunistic access," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4040–4050, Sep. 2009.
[27] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar, "Providing quality of service over a shared wireless link," *IEEE Commun. Mag.*, vol. 39, no. 2, pp. 150–154, Feb. 2001.
[28] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1277–1294, Jun. 2002.

Shang-Pin Sheng received the BEng degree in electrical engineering in 2008 from National Taiwan University, Taipei, Taiwan, the MEng degree, and the PhD degree in electrical engineering from the University of Michigan, Ann Arbor, in 2011 and 2014, respectively. His research interests include dynamic spectrum access, network economics and pricing.

Mingyan Liu (M-'00, SM-'11, F-'14) received the BSc degree in electrical engineering, in 1995, from the Nanjing University of Aero. and Astro., Nanjing, China, the MSc degree in systems engineering, and the PhD degree in electrical engineering from the University of Maryland, College Park, in 1997 and 2000, respectively. She joined the Department of Electrical Engineering and Computer Science at the University of Michigan, Ann Arbor, in September 2000, where she is currently a professor. Her research interests include optimal resource allocation, performance modeling and analysis, and energy efficient design of wireless, mobile ad hoc, and sensor networks. She is the recipient of the 2002 US NSF CAREER Award, the University of Michigan Elizabeth C. Crosby Research Award in 2003, and the 2010 EECS Department Outstanding Achievement Award. She serves/has served on the editorial board of the *IEEE/ACM Transactions on Networking*, *IEEE Transactions on Mobile Computing*, and *ACM Transactions on Sensor Networks*. She is a fellow of the IEEE.

Romesh Saigal received the BTech (Hons) degree in mechanical engineering in 1961, the MTech degree in Industrial engineering, in 1963, from the Indian Institute of Technology, Kharagpur, India., and the PhD degree in industrial engineering and operations research from the University of California, Berkeley, in 1968. He subsequently taught at the Schools of Business at Berkeley (1967-1973), Northwestern University, Evanston, Ill, from 1976 to 1986, and worked as a member of technical staff at the Bell Telephone Laboratories, Holmdel, NJ, from 1973 to 1976. He joined the Department of Industrial and Operations Engineering at the University of Michigan, Ann Arbor, in September 1986, where he is currently a professor. His teaching and research interests are in optimization and financial engineering and he has served as a director of the Financial Engineering Program at the University of Michigan, Ann Arbor, MI. He has served on the editorial board of *Management Science*.

Queries to the Author

Q1. Index term(s) in the article is missing. Please provide the Index term(s).
Q2. Please provide zip codes.
Q3. Please provide page range for Ref. [2].
Q4. Please provide publisher location for Ref. [23].
Q5. Please provide the photograph of all authors.