I. INTRODUCTION

The use of efficient and clean renewable energy sources is one of the major conditions required to achieve the important aim of sustainable development in modern societies [1]. Wind energy is one such source and wind turbines are being subject to intensive studies for improving their efficiency [2,3]. Although the basic laws of atmospheric wind motion have been known for a long time, important problems such as turbulence, layering, and the statistics of extreme events remain poorly understood. A better understanding of these phenomena can help in the construction of energy conversion schemes that are both more efficient and robust. Here two aspects of robustness must be considered. First, the occurrence of sudden changes in wind speed and direction can interrupt the process of energy conversion, meaning unreliability and a sudden slump in the electrical energy generated, which is seen as one of the major obstacles for the replacement of fossil and nuclear plants by wind energy sources. Second, these sudden changes introduce massive mechanical stresses that can lead to excessive wear or, ultimately, to the destruction of wind generators.

Wind flow is in general turbulent [4] and nonhomogeneous [5] with a non-negligible stochastic contribution. Therefore, in order to be able to construct more accurate models for its physical properties, one needs either accurate measurements of the wind speed on the length scales of wind turbines or suitable models that can statistically reproduce these measured data. Since the wind turbines are driven by turbulent wind fields, the stochasticity of the wind fields transfers to the stochastic dynamics of the wind turbine as a whole, of the loads on its structures, and of the power output. Recently, a Markov chain model was used to reproduce wind measured data [6], based on the transition matrix and time propagators for the wind speed and direction together with the power production. Differently from previous first-order approaches [7], information from two- and three-step transition probabilities are considered.

In this paper we aim at understanding the stochastic aspects of power production coupled to the wind velocity field. To that end, we use a methodology introduced in Ref. [8], for uncovering optimal stochastic variables weakly [9] and strongly coupled [8], and adapt it with two purposes: first, to properly derive the functional relation of pairs of variables whose values are extracted from the Markov chain model for wind turbines and second, to uncover specific features of the wind turbine and characterize the different working regions observed in the power-speed plane. After this, we test our approach to uncover the functional dependence of the well-known performance curve, which describes the functional dependence of the power production and the wind speed. Whereas previous reports have pointed out the benefits of deriving the power curve from the drift field [10], we additionally take the diffusion field into account and find that this procedure creates additional insight.

We use the data sets generated by the Markov chain model described by Lopes et al. [6]. Using such synthetic data sets that properly reproduce the statistical features of empirical data sets allows us to use data sets as large as needed for our analysis. Moreover, the Markov chain model serves as a filter to remove periodicities present in the data. This is an essential step for our modeling, as it would not be possible to correctly estimate the Kramers-Moyal coefficients from the raw data.

We start in Sec. II by describing the empirical data used to define the Markov chain model as well as the data generated with it. In Sec. III we describe our stochastic method for analyzing the data and in Sec. IV we apply it to analyze the performance curve of a wind turbine. In Sec. V we show that the same method, when applied separately to both wind speed and power production, allows us to derive the performance curve. Further, the same analysis also provides insight concerning specific features of the turbine system studied. Section VI presents a discussion and conclusions.

II. PROPERTIES AND GENERATION OF THE DATA SETS

The data analyzed in this paper were simulated from a set of measurements from a wind turbine in the region of
lead to economic benefits from scheduling and maintaining a level of constant production.

To overcome some of these problems, we employ a reconstruction of the original wind data set through a Markov Chain Model (MCM), which has recently been established [6] using a joint discretization of the wind speed, power, and direction variables for the state definition. In our specific case within the range of each variable \( P, v, \) and \( \theta \) we select 80, 60, and 12 states. States without a realization in the time series are deleted and all states that contain at least one observation are kept. Details about the estimation of the transition probabilities can be found in Refs. [6,16] and Appendix A. The use of the MCM has the additional benefit of removing periodicities from the data.

Based on the Markov chain transition matrix \( P \) with \( P(i,j) = p_{i,j} \) being the probability of transition from state \( s_i \) to state \( s_j \), the synthetic data sets were generated using the following Monte Carlo approach. We find the cumulative probability transition matrix \( P_{\text{cum}} \) with \( P_{\text{cum}}(i,j) = \sum_{k=1}^{i-1} p_{i,k} \) and select randomly an initial state \( s_i \). A random number \( \epsilon \) between zero and one is then uniformly selected and a new state \( s_N \) is chosen such that \( P_{\text{cum}}(s_i,N) > \epsilon \). (For details see Ref. [7].) Figure 2 shows the generated time series. As shown in Ref. [6], the proposed Markov chain model reproduces the dominant statistical features of all three properties, namely, power production, wind speed, and wind direction, although no periodicities are present in the reproduced time series.

The resulting synthetic data series for power production and wind speed also retain the persistence statistics, namely, the average duration of power production and wind speed on a certain level, respectively. Power production and wind speed are presented as fractions of the maximum observed power \( P_{\text{max}} \) and wind speed \( v_{\text{max}} \), respectively, assuming, therefore, values between zero and one.

The wind direction in the original time series follows a bimodal pattern, which is, to some extent, also periodic: During the day there are weaker breezes in a particular direction than during the night when wind streams in a different direction; therefore, the prevailing wind speed values occurring during the "day" (2 p.m. to 2 a.m.) have approximately \( \theta_1 \sim 80^\circ \), a value different from the one found for wind speeds measured during the "night" (2 a.m. to 2 p.m.), \( \theta_2 \sim 320^\circ \). The Markov chain model is capable of preserving this bimodality. Although we will not consider the wind direction in our study, one should notice that the wind direction bimodality is reflected in the distribution of the wind speed (see the inset of Fig. 5). Another bimodality preserved from the original data set is the one in the joint probability density function (PDF) of velocity and power, which is visible in Fig. 4(f).

Figure 3 shows the discretization of the data set where the combined states for power and speed are indicated with boxes and circles represent the historical data set points projected into the speed-power plane. The synthetic data sets were generated with \( 2 \times 10^6 \) data points and show stationary behavior, i.e., have constant moving averages (not shown).

This approach has several advantages over the direct analysis of historical data sets. First, high-quality data series of arbitrary length can be generated, which increases the accuracy of the Markov analysis. Second, the generated data are by construction Markovian, with the reconstruction through the
FIG. 2. (Color online) Time series for (a) the magnitude of the wind speed \( v \). The inset shows the time series for a shorter time period. The PDF of the wind speed is shown in the inset of Fig. 5. (b) Power production \( P \) of the wind turbine. The inset shows the PDF of the time series. (c) Corresponding wind direction \( \theta \). The inset shows the PDF of the time series. All data series were generated with the Markov chain model [6] described in Sec. II. All properties are normalized to the observed intervals \([0, v_{\text{max}}], [0, P_{\text{max}}], \) and \([0, \theta_{\text{max}}] \), respectively. In this and all following figures, multiples of 10 min, whereas velocity \( v \), power \( P \), and direction \( \theta \) are normalized to unity.

Markov chain acting as a filter that removes both noise correlations and periodicities. Finally, non-Gaussian transition probabilities between the states are preserved, which enables us to study them through higher Kramers-Moyal coefficients. Error analysis for the derivation of the transition matrix is described in Appendix B.

III. STOCHASTIC ANALYSIS OF WIND TURBINES

The coevolution of two or more stochastic variables, such as wind speed and power production, can be described through a system of coupled stochastic equations, each one defined by a deterministic contribution (drift) and stochastic fluctuations from possible stochastic sources. In this section we present the general framework to analyze our data and in the next section we apply it to the power production and wind speed variables.

For the general case of \( K \) stochastic variables \( X_1, \ldots, X_K \), the vector \( \mathbf{X}(t) = [X_1(t), \ldots, X_K(t)] \) defines the state of the system under study at each time instant \( t \). The evolution of the state vector yields a stochastic trajectory in phase space and is given by the so-called Itô-Langevin equation [8,14,17]

\[
\frac{d\mathbf{X}}{dt} = \mathbf{h}(\mathbf{X}) + \mathbf{g}(\mathbf{X})\mathbf{\Gamma}(t),
\]

where \( \mathbf{\Gamma}(t) = [\Gamma_1(t), \ldots, \Gamma_K(t)] \) is a set of \( K \) independent stochastic forces with Gaussian distribution fulfilling the following conditions: \( \langle \Gamma_i(t) \rangle = 0 \) and \( \langle \Gamma_i(t) \Gamma_j(t') \rangle = 2\delta_{ij}\delta(t-t') \). The function \( \mathbf{h} = \{h_i\} \) in Eq. (1) is the deterministic contribution, describing the physical forces that drive the system, while \( \mathbf{g} = \{g_{ij}\} \) describes the amplitude of the stochastic sources of fluctuations \( \mathbf{\Gamma} \) [18].

The evolution of the stochastic variables in time yields a joint PDF \( f(\mathbf{X}) \) that evolves according to the so-called Fokker-Planck equation

\[
\frac{\partial f(\mathbf{X},t)}{\partial t} = -\sum_{i=1}^{N} \frac{\partial}{\partial X_i} \left[D_i^{(1)}(\mathbf{X})f(\mathbf{X},t)\right] + \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\partial^2}{\partial X_i \partial X_j} \left[D_{ij}^{(2)}(\mathbf{X})f(\mathbf{X},t)\right].
\]
where the functions $D_i^{(1)}$ and $D_{ij}^{(2)}$ are related to the functions $h_i$ and $g_{ij}$ above, namely,

$$D_i^{(1)}(X) = h_i(X),$$

$$D_{ij}^{(2)}(X) = \sum_{k=1}^{N} g_{ik}(X)g_{jk}(X),$$

and are usually called drift and diffusion functions, respectively. Drift and diffusion functions can be directly derived from observed or generated data [18,19] and this fact is the basis of our framework. Indeed, the drift and diffusion functions of the underlying process are defined through conditional moments, namely [14],

$$D^{(k)}(X) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \langle M^{(k)}(X,\Delta t) \rangle,$$

where $M^{(k)}$ are the first and second conditional moments ($k = 1, 2$). These conditional moments can be directly derived from the measured data as [18,19]

$$M_i^{(1)}(X,\Delta t) = \langle Y_i(t + \Delta t) - Y_i(t) | Y(t) = X \rangle$$

and

$$M_{ij}^{(2)}(X,\Delta t) = \langle [Y_i(t + \Delta t) - Y_i(t)][Y_j(t + \Delta t) - Y_j(t)] | Y(t) = X \rangle,$$

where $Y(t) = [Y_1(t), \ldots, Y_N(t)]$ is the $N$-dimensional vector of measured variables and $\langle \cdot | Y(t) = X \rangle$ symbolizes a conditional averaging over the entire measurement period, where only measurements with $Y(t) = X$ are taken into account. Important conditions to hold are that (i) the underlying process is stationary and (ii) the Markovian property is fulfilled.

Numerically, $h$ and $g$ are determined on an $n_1 \times \cdots \times n_N$ mesh of points in phase space, as a function of the variables $X_i$, using the drift and diffusion functions. Locally, at each mesh point, one can always diagonalize the matrix $g(X)$ and compute their $K$ eigenvalues and $K$ eigenvectors. As shown previously [8,15,20], this analysis provides information about the stochastic forces acting on the system. Namely, the eigenvalues indicate the amplitude of the stochastic force and the corresponding eigenvector indicates the direction toward which such force acts. In a previous work [8] we argued that to each eigenvector of the diffusion matrix one can associate one independent source of stochastic forcing $\Gamma_i$ and thus the eigenvectors can be regarded as defining principal axes for stochastic dynamics. In particular, if one eigenvalue is very small compared to all the others, the corresponding stochastic force can be neglected. In the following sections we present a different implication of this principal stochastic component analysis, which emphasizes that the vanishing of one stochastic direction is in fact an indication of a strong functional dependence between the pair of variables being analyzed.

### IV. WIND TURBINE DRIFT AND DIFFUSION MAP ANALYSIS

In this section we focus solely on two variables: power production $P$ and wind speed $v$. Since both series are stationary and Markovian, we assume them to evolve according to the following equations:

$$\frac{dv}{dt} = h_v(v, P) + g_{vv}(v, P)\Gamma_1 + g_{vP}(v, P)\Gamma_2,$$

$$\frac{dP}{dt} = h_p(v, P) + g_{pp}(v, P)\Gamma_1 + g_{pP}(v, P)\Gamma_2.$$

In general, the six functions defining the vector $h$ and matrix $g$ depend on both variables and describe the coupling between each other. Based on Eqs. (3), we can derive both $h$ and $g$ from the functions $D^{(1)}$ and $D^{(2)}$, which in turn are extracted directly from the synthetic data set by computing the corresponding conditional moments using Eq. (4). Note that solving Eq. (3b) for computing the matrix $g$ yields multiple solutions. If $g$ is a solution then all matrices of the form $\tilde{g} = gO$, where $O$ is an orthogonal matrix ($OO^T = I$), are also admissible solutions. The matrix $g$ can therefore be computed as the square root of matrix $D^{(2)}$, i.e., by diagonalizing $D^{(2)}$ through a proper permutation matrix $P$, and since all eigenvalues are positive ($D^{(2)}$ is positive definite), taking the square root of each eigenvalue and transforming the matrix back.

Figures 4(a)–4(e) show the five components of $D^{(1)}$ and $D^{(2)}$, i.e., the numeric results for both the drift and the diffusion coefficients computed directly from the generated $P$ and $v$ time series. The large fluctuations in the region near maximum power production and wind speed are due to a lack of observations. Indeed, the joint PDF for $P$ and $v$ [Fig. 4(f)] shows that this region is poorly sampled.

To extract valuable information, next we treat these functions separately. Namely, we consider the drift vector field $(h_v, h_P)$ and the eigenvectors of the diffusion matrix associated with its eigenvalues $\lambda_{\text{max}}$ and $\lambda_{\text{min}}$. Figure 5 shows the drift vector field in the power production and wind speed state space, restricted to the sampled region defined by the power production curve in Fig. 3. The solid black line is the performance curve computed from the $(P, v)$ joint probability density function, shown in Fig. 4(f), and defines the most likely power production for a given wind speed. Three different regions can be identified.

Region I in Fig. 5 is characterized by a high wind speed, i.e., above a threshold velocity $v_{\text{th}}$ that exceeds the rated wind speed of the turbine. Postponing a more detailed description to Sec. V, we define $v_{\text{th}} = 0.55v_{\text{max}}$, which corresponds to the 97th percentile of the wind speed distribution. For these rare events of high wind speeds, the expected behavior of the wind turbine is to maintain the power production since there is a surplus of energy in the airflow. In this region, the performance curve is roughly constant at $\sim 0.95P_{\text{max}}$. Still, positive power drifts are observed whenever the power production is below the performance curve. The wind speed drift is large in magnitude and always negative, i.e., the drift points towards lower wind velocities.

Region II is characterized by production levels above a transition region of $0.6P_{\text{max}} \lesssim P \lesssim 0.8P_{\text{max}}$, indicated by gray shading in Figs. 5 and 6, which is defined in more detail in Sec. V and Fig. 6. A closer look at region II enables one to identify a fixed point region ($D^{(1)} \sim 0$) at high power production levels and wind speed $v \sim 0.5v_{\text{max}}$, shown as an encircled area in Fig. 5 near the $v_{\text{th}}$. This speed value coincides approximately with the rated wind speed, i.e., the speed for...
which the turbine was designed and at which it operates at an optimal regime. It can therefore be concluded from our analysis that the turbine has been well selected and it remains to be seen if similar conclusions can be drawn when applying our method to arrays of turbines.

Finally, region III is characterized by frequent low-speed events with a power production below \( P \approx 0.6P_{\text{max}} \), containing another attraction point at \( v_{\text{mode}} \approx 0.14v_{\text{max}} \).

In previous works [5] the drift vector field around the performance curve was parallel to the power production axis. In Fig. 5 the vector field tends to be tilted towards the performance curve, because the data analyzed was sampled with a much smaller frequency, and therefore the time between successive measures is sufficiently large to observe the convergence to the stable fixed points.

Such observations can be more clearly understood by considering Fig. 5 together with the marginal PDF of the wind speed shown in its inset. The distribution of observed values for the wind speed follows approximately a Weibull distribution, as is known from the literature [21]. In our case we observe significant deviations at the tail, which shows a bump. This deviation can be explained with the help of Fig. 2(c) and by recalling that there are correlations between the wind speed with each one of the main wind directions (as highlighted in Sec. II), resulting in the bimodality of the wind speed. The bump in the distribution for the wind
FIG. 5. (Color online) Drift vector $D^{(1)}(P,v) = [h_P(P,v), h_v(P,v)]$ [see Eq. (5)] in each $(P,v)$ box used to generate the data (see Fig. 3). Three regions can be identified: region I having slow dynamics and regions II and III with fast dynamics. Interestingly, the fixed point spot in region II coincides with the rated speed of the wind turbine. The black curve indicates the performance curve and crosses the bins for which $D^{(1)}(v,P)$ vanishes. In the inset, the marginal probability density function of the wind speed $v$ is well fitted by a Weibull distribution with scale parameter $\lambda \approx 0.25$ and shape parameter $k \approx 1.66$ and a mode $v_{\text{mode}} \approx 0.14 v_{\text{max}}$ [21]. Velocities above $v_{\text{th}} = 0.55 v_{\text{max}}$ are rarely observed, which explain the observed drifts in region I. The gray shading indicates the gradual transition from region II to region III. Dotted vertical lines mark the positions of $v_{\text{mode}}$, $\langle v \rangle$, and $v_{\text{th}}$, respectively.

speed indicates one of the two modes, namely, the one observed at high wind speeds. Therefore, the bimodality of the original data is preserved as bimodality of the Markov chain model.

The two dominant trends identified in region I are compatible with the expected behavior of the power production control system present on the wind turbine. For high values of wind speed, the controller action upon the blade aerodynamics is capable of sustaining the production level despite the expected decrease of the wind speed. One has to consider, however, the time scales involved. With a 10-min resolution of the original data, it is not possible to directly observe the rapid controller action on the blades, only the average behavior of the controller as well as actions that occur on larger time scales, such as the rotation of the tower. However, even at large sampling times, the data set catches some events beyond the power curve and the subsequent conditional moments mirror the controller action that forces the system back on the curve.

Another important application of our method deals with the diffusion matrix. As explained in the previous section, by diagonalizing the diffusion matrix at each point of the phase space one is able to determine the two eigendirections for diffusion. Being orthogonal to each other, these two eigendirections define an ellipse with major and minor axes proportional to the corresponding eigenvalue. Figure 6 shows the diffusion ellipses in phase space. Region I is characterized by the largest ellipses indicating very large fluctuations, while there is an area in region II that presents small fluctuations and corresponds to the fixed point areas identified in the drift field. In the high-slope region of the power curve, the ellipses degenerate, i.e., one eigenvalue is negligible when compared to the other ($\lambda_{\text{min}}/\lambda_{\text{max}} \sim 0$). The inset of Fig. 6 shows in a grayscale the quotient $\lambda_{\text{min}}/\lambda_{\text{max}}$ between the smallest and the largest eigenvalues. White corresponds to zero quotient, while values in [0.2,1] are black. Clearly, a white region indicating a very low ratio of the eigenvalues can be identified, which follows the performance curve shown in Fig. 3. Details concerning our error analysis are described in Appendix B.

One remark is appropriate at this point. As we mention above, the procedure described in this section has been performed on synthetic data generated using a Markov chain model. One might argue that the drift and diffusion coefficients could be extracted directly from the measured data. However, as a process in time, the real data also reflect daily and seasonal variations, which hinders and eventually spoils this approach. It has been found (Sec. IV) that the dynamics of
the system are reflected by the presence of two maxima in the joint distribution (one at low \( v \), near \( v_{\text{mode}} \) and one at high \( v \), near \( v_{\text{th}} \)). The two-dimensional PDF of the measured data (not shown) closely resembles the one of synthetic data, shown in Fig. 4(f), including the two maxima. Estimating drift and diffusion coefficients directly from the measured data, however, does not allow us to reproduce this distribution (not shown). The MCM has been found to reproduce the relevant statistical and dynamical features of the fluctuations observed in the real data [6]. Given that (i) the MCM produces a transition matrix that maximizes the likelihood of distributions [6], (ii) its errors are known and small (see Appendix B), (iii) it faithfully reproduces the joint PDF, and (iv) the analysis using the MCM yields the correct physical dynamics and fixed points (see Sec. IV), it is reasonable to take the simulated data as the aperiodic process corresponding to the time series of power and wind speed.

V. DERIVING THE PERFORMANCE CURVE FROM UNIVARIATE STOCHASTIC DYNAMICS

The inset of Fig. 6 shows that along the performance curve one eigenvalue is typically much larger than the other. This is an indication that in fact \( P \) is a function of \( v \), which in the case of power production and wind speed yields the performance curve drawn in Figs. 3, 5, and 6. To see this one first takes \( P \) and \( v \) as two general variables fulfilling Eqs. (5) and observes that if \( P \equiv P(v) \) there are not two independent stochastic forces, but only one, yielding for \( v \) [17] and for \( P \)

\[
\begin{align*}
\frac{dv}{dt} &= \tilde{h}_v(v) + \tilde{g}_v(v)\Gamma', \\
\frac{dP}{dt} &= \tilde{h}_P(P) + \tilde{g}_P(P)\Gamma',
\end{align*}
\]
where functions $\tilde{h}$ and $\tilde{g}$ are of course different from the drift and diffusion functions defined above in Eqs. (5), since only one variable is taken into consideration for the stochastic motion equation.

Consequently, the separate analysis of both wind speed and power production enables one to extract valuable insight about the full dynamics and behavior of the wind turbine with the atmospheric wind. In fact, the two-dimensional analysis of the performance curve summarized in Figs. 5 and 6 can indeed be accessed through a one-dimensional stochastic analysis of each variable $P$ and $v$ separately.

Figure 7 shows the drift and diffusion of both the wind speed and power production determined for the model described by Eqs. (6a) and (6b). The drift of the wind speed $\tilde{h}_v$ has three zeros. These zeros correspond to three fixed points, two stable ($v \approx v_{\text{mode}} \approx 0.14v_{\text{max}}$ and $v \approx v_{\text{th}} \approx 0.55v_{\text{max}}$) and one unstable at $v_{\text{UFP}} \approx 0.42v_{\text{max}}$. Thus, for wind speed below $v_{\text{UFP}}$, the airflow is unstable and unsuited for power production, while wind speeds above $v_{\text{UFP}}$ promote power production. The first zero of $\tilde{h}_v$ indicates approximately the mode of the wind speed distribution (compare with the inset in Fig. 5) and the other two zeros mark the transition between two different regions identified above in Figs. 5 and 6. The transition between regions I and II is marked by $v_{\text{th}}$. The transition between regions II and III is more subtle and deals with the zero at $v_{\text{UFP}}$ and with the functional dependence of the two variables: It is located at the transition region of $0.6P_{\text{max}} \lesssim P \lesssim 0.8P_{\text{max}}$, which corresponds to $v_{\text{UFP}} \approx 0.42v_{\text{max}}$ (cf. Fig. 5). The transition between regions III and II is also located at a minimum of the diffusion $\tilde{g}_v$ for the wind velocity. Moreover, Fig. 7(a) also shows that positive drifts are located at small wind speeds (up to $v_{\text{mode}}$) and for region II (see Figs. 5 and 6). At $v_{\text{UFP}}$, the drift changes to a positive value. In other words, above $v = v_{\text{UFP}}$ the expected change of the wind speed is towards higher values. The drift reaches a maximum on region II prior to a steep change towards negative values. In this region, the wind speed values are not in the range of the extreme weather conditions and are also not as frequent as the lower wind speed values. However, it is frequent enough to be associated with a commonly repeated pattern, i.e., the bimodal pattern of the wind direction, in which the airflow is mainly induced by thermal differences. This pattern is responsible for most of the power production in this wind turbine and the main reason for a second attraction point at $v_{\text{th}}$. For higher wind speed values, the drift changes again to negative values (region I) since very high wind speed is usually of short duration, i.e., extreme wind gusts. Moreover, Fig. 7(b) shows where the power stochasticity is maximal, i.e., the region of the highest diffusion values $\tilde{g}_P$, which is the high-slope region of the power curve for $0.3P_{\text{max}} \lesssim P \lesssim 0.6P_{\text{max}}$.

Parallel to our stochastic approach, we also derive coefficients $D^{(1)}$ and $D^{(2)}$ directly from the transition matrix. The results are similar to the ones described above, as one can see in Fig. 7 (thin gray solid and dashed lines), where the deviations from the coefficients derived from our analysis are due to the prescribed binning of phase space, i.e., to the number of states chosen for the transition matrix. Taking all the above observations into account, one concludes that there is strong agreement between the regions defined in the context of Figs. 5 and 6 and the sign of the wind speed drift.

Having analyzed separately both properties $v$ and $P$, we continue by showing that from the drift and diffusion coefficients $\tilde{h}_v$, $\tilde{g}_v$, $\tilde{h}_P$, and $\tilde{g}_P$, one obtains a functional dependence between power production and wind speed. To that end we assume that Eq. (6a) holds for $v$ and that the other variable $P$ is an exclusive function $P(v)$ of $v$. Thus we can take the Itô-Taylor expansion [17] of its differential

$$
\begin{align*}
    dP(v) &= P(v + dv) - P(v) \\
    &= \frac{dP}{dv} dv + \frac{1}{2} \frac{d^2P}{dv^2} dv^2 + \mathcal{O}(dv^3) \\
    &= \left( \frac{dP}{dv} \tilde{h}_v + \frac{1}{2} \frac{d^2P}{dv^2} \tilde{g}_v^2 \right) dt + \frac{dP}{dv} \tilde{g}_v dw
\end{align*}
$$

(7)

using the differential $dv = \tilde{h}_v(v) dt + \tilde{g}_v(v) dw$. Therefore, identifying

$$
\begin{align*}
    \tilde{h}_P &= \frac{dP}{dv} \tilde{h}_v + \frac{1}{2} \frac{d^2P}{dv^2} \tilde{g}_v^2, \\
    \tilde{g}_P &= \frac{dP}{dv} \tilde{g}_v.
\end{align*}
$$

(8a)  (8b)
which can be solved with respect to the two derivatives of $P(v)$, yields the numerical integration scheme as follows:

$$P(v + \Delta t) = P(v) + \frac{dP}{dv}_{v,P(v)} \Delta v + \frac{1}{2} \frac{d^2 P}{d^2 v}_{v,P(v)} (\Delta v)^2 + O((\Delta v)^3)$$

$$= P(v) + \frac{\tilde{g}_P(P(v))}{\tilde{g}_v(v)} \Delta v + \frac{1}{2} \frac{\tilde{h}_P(P(v))}{\tilde{g}_v(v)^3} (\Delta v)^2 + O((\Delta v)^3).$$

(9)

Figure 8 shows the integration of $dP(v)$ for the condition $P_0(v_0) = 0$ for $v_0 = 0$. The deviations can be attributed to the fact that Eqs. (6a) and (6b) are strictly only valid in the regions where the eigenvalues of the diffusion matrix show a large difference between them, $\lambda_{\text{min}} \ll \lambda_{\text{max}}$ (cf. the insets of Figs. 5 and 6 where 0.2 $\ll v \ll v_{\text{th}}$). In addition, the integration Eq. (9) is valid only on the performance curve and applying Eqs. (6a) and (6b) therefore also neglects the asymmetry of the drift functions with respect to this curve. Both deviations are a natural consequence of having treated two dependent variables $P$ and $v$ separately.

VI. DISCUSSION AND CONCLUSIONS

Investigating a wind turbine from a real wind park, we report the reconstruction of the stochastic performance curve of the variables of wind speed and power production, using both drift and diffusion coefficients. These coefficients, describing the respective deterministic and stochastic interactions of wind field, turbine aerodynamics, and controller action, are estimated from a synthetic time series generated using a Markov chain model of the original measurement data. We argue that this reconstruction is superior to a direct evaluation of the measurements.

As a main finding we present the fact that the reconstruction of the power curve using both drift and diffusion coefficients uncovers additional information not visible in an analysis of the drift field alone [10], even though we are using measured data of a very low measurement rate as model input. Specifically, our analysis reveals the existence of various distinct regions in the wind speed-power production plane. In addition, we have been able to reconstruct the power curve from the drift and diffusion coefficients using a method that should be able to uncover functional relationships between stochastic variables in a wide range of experimental setups. It should be noted that it is possible to infer the correlation between $v$ and $P$ from the joint PDF in Fig. 4(f) alone; however, such an approach neglects the dynamical behavior of the system.

Without our approach summarized in Figs. 5–7 one could not so surely claim the existence of three separate regions. Moreover, attached to these three regions we detected three fixed points of the dynamics, two of them stable and one unstable. These three fixed points are only shown clearly in Fig. 7, after performing the one-dimensional analysis with our method. In particular, considering the upper stable fixed point, while its coincidence with peak production can be identified directly in the $(P,v)$ joint PDF, the analysis sketched in Fig. 7 allows us to detect the region that belongs to its basin of attraction, within which drift drives the trajectories to the upper stable fixed point, whereas beyond this region trajectories are pushed to the lower stable fixed point.

Finally, information about how trajectories diffuse in phase space, i.e., the entire dynamics of the system, can only be obtained completely after extracting the drift and diffusion fields together with the fixed points of the drift field and the principal directions of diffusion. Only after analyzing the separation of diffusion eigenvalues visible in Fig. 6 can we postulate the existence of a single diffusive force underlying Eqs. (6a)–(6c).

Although there are limitations in increasing the extracted power of a wind turbine, typically described by the power coefficient that has a maximum value given by the Betz limit [22], our approach may be helpful in obtaining a better understanding of the complex dynamics that determines power production in wind turbines. Indeed, we believe that such a stochastic description, if applied to an entire wind park, would enable one to better quantify the risk associated with the estimate of global energy production. The global energy production of a wind park is usually determined by financial constraints, i.e., by decision making of how much energy one must buy or sell in the market to compensate for the energy production fluctuations. Therefore, we are now extending this methodology in order to consider coupled systems of wind turbines in nearby locations, using direct measures of power and wind speed as well as simulated data.

ACKNOWLEDGMENTS

The authors thank Matthias Wächter, David Klein- hans, and Maria Haase for useful discussions and GEN- ERG, SA for providing the original data. The authors
acknowledge partial support under Grants No. PEst-OE/FIS/UI0618/2011 and No. FCOMP-01-0124-FEDER-016080. Also, F.R. (Grant No. SFRH/BDP/65427/2009), T.S. (Grant No. SFRH/BD/86934/2012), and P.G.L. (Grant No. Ciência 2007) thank Fundação para a Ciência e a Tecnologia (FCT) for financial support. This work was part of a bilateral cooperation DRI/DAAD/1208/2013 supported by FCT and Deutscher Akademischer Auslandsdienst.

**APPENDIX A: ESTIMATION OF THE MARKOV CHAIN TRANSITION MATRIX**

The Markov chain transition probabilities were not obtained by employing the usual maximum likelihood (ML) estimator, but by using a modified likelihood function that combines information from one-step and two-step transitions, as described in this appendix. This modified ML yields lower variance estimates for the transition probabilities (see Appendix B).

Let \( \{s_i | i \in [0, \ldots, n]\} \), \( n \in \mathbb{N} \), be the state space of the Markov chain model, where each state is a combination of a wind power, speed, and direction state. Thus the wind power, speed, and direction time series can be transferred into a stream \( s \) of states, i.e., \( s = \{s_0, s_1, s_2, \ldots, s_{i-1}, s_i, s_{i+1}, \ldots, s_{m-1}, s_m\} \), where \( m \in \mathbb{N} \) denotes the length of the time series and \( s_i \) denotes the state the Markov process assumes at time \( k \) with \( k \in \{0, \ldots, n\} \). Let \( p_{ij} \) denote the probability of the process to move from state \( s_i \) at time \( k \), to state \( s_j \) at time \( k+1 \). Then the likelihood function \( L^{(1)} \), i.e., the probability of the observed series of states \( s \), given the transition probabilities \( p_{ij} \) for one-step transitions, is

\[
L^{(1)} = P(s_0)p_{i_0,i_1} \cdots p_{i_{n-2},i_{n-1}}p_{i_{n-1},i_n}.
\]  

(A1)

Since the probability of the process being in state \( s_i \) is constant, the maximum likelihood estimator can thus be written as the maximum of \( L^{(1)} = \prod_{(i,j) \in S_1} p_{ij} \) subjected to \( p_{ij} \geq 0 \) and \( \sum_{j=0}^n p_{ij} = 1 \), with \( i, j = 1, \ldots, n \), where \( S_1 \) is the set of all one-step transitions \((i, j)\) observed in the state stream \( s \).

Solving the above optimization problem is equivalent to minimizing the negative log-likelihood function \( L^{(1)}_{\text{log}} = -\sum_{(i,j) \in S_1} \log p_{ij} \), subjected to the same constraints. For practical purposes we consider henceforth the log-likelihood function \( L^{(1)} \).

Next we consider only two-step transitions. Taking \( p'_{ij} \) as the probability of the process moving from state \( s_i \) at time \( k \), to state \( s_j \), at time \( k+2 \), the maximum likelihood estimators are given by the maximum of \( L^{(2)} = \prod_{(i,j) \in S_2} p'_{ij} = \prod_{(i,j) \in S_2} \sum_{k=1}^n p_{ik}p_{kj} \), with \( S_2 = S_2^{\text{odd}} \cup S_2^{\text{even}} \), where \( S_2^{\text{odd}} = \{(i_1, i_3), (i_3, i_5), \ldots, (i_{m-3}, i_{m-1})\} \) (odd time points) and \( S_2^{\text{even}} = \{(i_0, i_2), (i_2, i_4), \ldots, (i_{m-2}, i_{m})\} \) (even time points), assuming without loss of generality that \( m \) is even. Here we use the Chapman-Kolmogorov equations to express the two-step transitions \( p'_{ij} \) in terms of one-step transitions \( p_{ij} \) with the same constraints as above. The corresponding log-likelihood function, whose minimization defines an equivalent optimization problem, is

\[
L^{(2)}_{\text{log}} = -\sum_{(i,j) \in S_2} \log(\sum_{k=1}^n p_{ik}p_{kj}).
\]

Finally, as described in Ref. [6], we minimize the objective function given by the sum of the two log-likelihood functions above, namely,

\[
L_{\text{log}} = L^{(1)}_{\text{log}} + L^{(2)}_{\text{log}}.
\]  

(A2)

and under the same constraints. Higher-order Markov models were also tested, but did not show considerable improvement compared to the estimator \( L_{\text{log}} \) in Eq. (A2). Notice that the minimization of \( L_{\text{log}} \) in Eq. (A2) for the series of states \( s \) is equivalent to minimizing the one-step functional \( L^{(1)}_{\text{log}} \) of the augmented state stream \( s' = s_{i_0}, s_{i_1}, s_{i_2}, \ldots, s_{i_{m-2}}, s_{i_{m-1}}, s_{i_m}, s_{i_1}, s_{i_2}, s_{i_3}, s_{i_4}, \ldots, s_{i_{m-2}}, s_{i_{m-1}}, s_{i_m}, s_{i_1}, s_{i_2}, \ldots, s_{i_{m-2}}, s_{i_{m}} \), where \( S \) stands for voids in the state stream.

Concerning the state space partition, we tried several different partitions of the power-speed-direction space. A comparison of the histograms of the original data and the data synthesized using these models showed that the model with the highest resolution of the binning in the power-speed plane is \( 80 \times 60 \times 12 \), which reproduces the original data most accurately.

**APPENDIX B: ERROR ANALYSIS**

A lower bound for the uncertainty of the transition probabilities is provided by the Cramer-Rao bound (CRB) since there is no simple expression to determine the parameters variance for the two-step estimator, described in Sec. VI. The CRB for maximum likelihood estimators with constraints on the parameter space can be computed by using the following equation [23]:

\[
\text{CRB} = U^T(U^H U^T)^{-1} U,
\]  

(B1)

where \( U \) is an orthonormal matrix spanning the null space of the Jacobian \( J \) of the parameter’s equality constraints and \( H \) is the Hessian of the objective function, defined in Appendix A.

The uncertainty is linked to the number of observed transitions in the data. Results show that a high value for the CRB (\( \sigma_{p_{ij}} > 0.1 \)) is always associated with infrequent transitions (fewer than four observations in the 2-yr data set). For transition probabilities \( p_{ij} \) with more than 25 observations, 90% of the \( \sigma_{p_{ij}} \) values are smaller than 0.25 \( p_{ij} \).

To assess the performance of the two-step estimator, its CRB is compared with the CRB of the one-step estimator, applied to the same 2-yr data set. Results shows that for the transition probabilities with a high CRB value, the two-step estimator provides a lower bound. In the remaining cases, the difference is not significant.

Another source of errors stems from our use of the direct estimation of the Kramers-Moyal (KM) coefficients from the synthetic time series. This method has been found to introduce three principal types of errors [18,24,25].

The first error accounts for the statistical variation of counts \( N \) in each bin used for the calculation of the conditional moments. It can be shown [25] that this error decreases with \( 1/\sqrt{N} \). Since we can generate synthetic time series of arbitrary length and therefore arbitrarily high \( N \) in each bin under consideration, we can neglect this error.

Second, estimation of the KM coefficients uses an expansion of the Fokker-Planck operator in powers of the temporal increment \( \tau \), neglecting higher orders in \( \tau \). It is known [24] that this finite-time expansion induces an erroneous count \( M^{(i)} \) of the \( i \)th conditional moment \( M^{(i)} \), namely, \( M_{i,\xi} = \tau D^{(i)} + \ldots \).
\[ \frac{1}{2} (D^{(1)} D^{(1)} + D^{(2)} D^{(2)}) + M_{2,E} = 2 \tau D^{(2)} + \tau^2 (D^{(1)} D^{(1)} + 2 D^{(2)} D^{(2)} + D^{(1)} D^{(2)} + D^{(2)} D^{(1)}). \]

We have calculated these errors numerically, using the estimated KM coefficients \(D^{(i)}\) and their numerical first and second derivatives \(D^{(i)}_1\) and \(D^{(i)}_2\), and found them to be generally within a few percent of the estimated KM coefficients.

A third source of error is the finite size of the bins used for the calculation of the conditional moments. Again, it can be shown that this finiteness induces an erroneous count of the moments (and corresponding coefficients)

\[ D_E^{(k)}(x_0) = \int_{x_0 - \Delta x}^{x_0 + \Delta x} D^{(k)}(x)p(x)dx, \]

where \(x_0\) and \(\Delta x\) are the respective bin centers and bin widths and \(p(x)\) is the stationary distribution of the stochastic variable. We numerically investigated this error using fits for both the stationary distributions and KM coefficients inside the bins and found it to be generally in the few-percent regime too, increasing considerably only at the edges of the regions investigated, where the KM functions become steeper.

The effect of the second and third error sources is indicated by error bars in Fig. 7.

**APPENDIX C: STOCHASTIC ANALYSIS OF TWO DEPENDENT VARIABLES**

In general, whenever the diffusion matrix \(D^{(2)}(P,v)\) has rank one, the set of variables \((P,v)\) has in fact only one independent stochastic source and therefore Eqs. (5) reduce to Eqs. (6), where the stochastic force \(\Gamma\) is the same for both variables. Consequently, one can write the differential of \(v\) as

\[ dv = \tilde{h}_v dt + \tilde{g}_v dW, \]

where the stochastic differential \(dW\) is the same as the one in Eq. (7). Since both variables \(P\) and \(v\) are driven by the same stochastic forces, one can take only one of them as the stochastic variable, say, \(v\), and the other one as a function of \(v\) and \(t\) alone, \(P = P(v,t)\). In that way one incorporates all stochastic contributions into \(v\). Mathematically this implies that we can write the differential of \(P\) as

\[ dP = \frac{\partial P}{\partial v} dv + \frac{\partial P}{\partial t} dt, \]

where both partial derivatives are derived from the functions \(\tilde{h}\) and \(\tilde{g}\) alone [see Eq. (6)]. Indeed, rewriting Eq. (7) as

\[ dP(v) = \frac{dP}{dv}(\tilde{h}_v dt + \tilde{g}_v dW) + \frac{1}{2} \frac{d^2 P}{dv^2} \tilde{g}_v^2 dt \]

and using Eq. (C1) together with the relations in Eq. (8) yields

\[ dP = \frac{\tilde{g}_v}{\tilde{g}_v} dv + \left( \tilde{h}_v - \tilde{h}_v \left. \frac{\tilde{g}_v}{\tilde{g}_v} \right| \right) dt. \]

This equation means that \(P\) is a function of \(t\) and \(v\), which contains all stochastic contributions. Consequently, the partial derivatives in Eq. (C2) are

\[ \frac{\partial P}{\partial v} = \tilde{g}_v, \]

\[ \frac{\partial P}{\partial t} = \tilde{h}_v - \tilde{h}_v \left. \frac{\tilde{g}_v}{\tilde{g}_v} \right| \]

The first partial derivative shows that the variation of the production power by speed variations equals the quotient of the corresponding diffusion amplitudes in time. The second partial derivative describes the local power production, which is given by the power production drift \(\tilde{h}_v\), after subtracting the contribution of the wind speed drift \(\tilde{h}_v\) in the variation of \(P\) due to \(v\).
