Detection and Statistics of Wind Power Ramps

Raffi Sevlian and Ram Rajagopal

***Abstract*—Ramps events are a significant source of uncertainty in wind power generation. Developing statistical models from historical data for wind power ramps is important for designing intelligent distribution and market mechanisms for a future electric grid. This requires robust detection schemes for identifying wind ramps in data. In this paper, we propose an optimal detection technique for identifying wind ramps for large time series. The technique relies on defining a family of scoring functions associated with any rule for defining ramps on an interval of the time series. A dynamic programming recursion is then used to find all such ramp events. Identified wind ramps are used to perform an extensive statistical analysis on the process, characterizing ramping duration and rates as well as other key features needed for developing future models.**

I. Introduction

Uncertainty in wind power production introduces significant cost in its integration to the grid [4], [5]. Efficient control strategies and market mechanisms can be used to mitigate this uncertainty [2], [15]. However, designing these mechanisms requires careful statistical models of wind power production [14].

Wind power has uncertainty in multiple time scales, each requiring its own modeling approach [7]. In the range of seconds to minutes, fast fluctuations of small amounts are modeled well with autoregressive or persistence techniques. These models however fail to capture longer term trends [11]. Autoregressive Integrated Moving Average (ARIMA) techniques, which include an integral step, can model such processes but do not capture abrupt changes in wind power generation [12]. On scales higher than a few minutes, trend modeling is needed.

A typical long term trend with large positive or negative change in a short period is called a ramp. Ramp event modeling offers an additional insight beyond ARIMA techniques by modeling these large deviations observed in wind power time series [6], [13] [3]. While autoregressive methods model the short term turbulence in the wind, ramp events can model more long term fluctuations which have a basis in the physical aspect of short term atmospheric dynamics. Up ramps are generally caused by intense low pressure systems, low level jets, thunderstorms, wind gusts or other longer term weather phenomena. Likewise, down ramps occur from the reduction or reversal of these physical processes [6]. These events are important in wind power management because large swings in power generation must be curtailed to prevent damage to wind turbines and power networks. Down ramps must be met with increasing generating capacity through other means. Developing statistical models offers insight into both forecasting of wind ramps as well as formulating stochastic control strategies for dealing with ramp events.

There exist two major problems in characterizing wind power ramps. The first is correctly detecting wind ramps in historical data. The detection problem requires precise definitions of wind ramps as well as a standardized technique for segmenting data into positive and negative ramp events which is robust to fluctuations in wind power in short time scales. Different organizations utilize different definitions of wind ramps. Given a set of definitions, current detection schemes rely on simplified and not always effective heuristic techniques. An optimal detection scheme which can incorporate any set of definitions allows broader sets of definitions to be explored and a systematic approach to ramp detection to be developed. The second important problem is modeling ramp events in a useful statistical framework. Statistical models offer insight into forecasting of wind ramps and into modifying power dispatch formulations to mitigate ramp impact. Building such models requires combining insight from physical processes with abundant data and an accurate detection algorithm for given definitions of a ramp event.

The contributions of this work are two fold. First we introduce an optimal ramp detection algorithm and discuss its implementation and approximations. We prove the optimality of our technique and then evaluate it under various conditions. Second we use publicly available data to perform optimal ramp detection and provide an in depth empirical statistical analysis of key ramp features. Such features can be used in future dispatch formulations to characterize ramp events.

The paper is organized as follows. Section II reviews previous work on the topic. Section III formulates the maximum weight subsequence problem to detect wind ramps. Section V presents results of experiments in detecting ramps for various datasets. Section VI introduces some simple statistical models to characterize ramp events. Section VII presents conclusions.

II. Previous Work

The field of wind ramp detection, characterization and modeling is a recent topic in the statistical forecasting community. Therefore, the definition of what constitutes wind ramp events has been arbitrary but has followed the general property of identifying large positive or negative swings in wind power output within a short time window. Another characteristic used in defining ramps has been high rate of power increase. There is no set definition on the minimum size of the increase or of the rate of increase. A consequence is that varying these definitions leads to different sets of detected ramp events.

There is also no clear detection method in the literature to segment and classify all wind power time series intervals that

R. Sevlian is with the Stanford Sustainable Systems Lab and Department of Electrical Engineering, Stanford University, USA. (e-mail: rsevlian@stanford.edu).

R. Rajagopal is with the Stanford Sustainable Systems Lab and Department of Civil and Environmental Engineering, Stanford University, USA. (e-mail: ramr@stanford.edu).

satisfy a given set of definitions in a reliable and reproducible way. Existing methods are unable of detecting all ramps of varying durations or rates which satisfy some minimum threshold value. This problem has been avoided generally by looking at large time scales (order of hours) and sufficiently smoothing the power series.

[8], [9] proposes a technique for detecting wind ramp events of a set length of 15, 30, and 60 minutes. Each time series sample is tested to see if it satisfies the ramp definition under each duration value. A scheme like this is adequate for large sample intervals in a sufficiently smoothed time series, but runs into a problem of counting large ramps multiple times. This work also provides a clear set of rules to test whether an interval of a wind power time series is in fact a ramp event or not.

Currently, the California Independent System Operator (CAISO) analyzes historical wind ramp data following the technique in [8]. An hour long window is tested to see whether the difference between maximum and minimum wind speed is greater than some threshold value. This test is performed every 5 minutes. Currently, CAISO only computes ramp swings statistics, and in particular the hourly distributions of ramp swing. Capacity reserve used by the system are characterized based on this metric. With a more accurate detection and classification scheme, more thorough modeling and analysis is possible.

Existing work on statistical modeling of wind ramps have focused on data driven approaches to predicting various quantities in wind ramp events, without addressing model formulation of the ramping process. [9] performs regression and feature selection on a binary variable indicating whether a particular day will correspond to a single ramp event or not. The results indicate that additional information acquired in nearby weather stations can be useful in predicting ramp events. The model only captures the time period of a day and additional work is needed for shorter time horizons. [17] uses various machine learning tools to predict ramp rates. The definition of a ramp rate is the first derivative of the underlying wind power time series. Similarly, in [16] ramps are grouped in classes and a support vector machine is proposed to predict and classify the next ramp event. These methods maybe successful in prediction, but their lack of structure makes them undesirable for model building for future control applications. In [1], a graphical model framework is proposes a switched ARMA model for wind. The model learns the inherent regimes of wind power in the data without incorporating ramp phenomena since regimes are hidden variables.

III. Problem Statement

A ramp represents a large increase or decrease in wind power within a limited time window. Ramps are parameterized by the following variables: ramp start, ramp duration and ramp rate. These uniquely identify such events and are shown in Figure 1. In addition to ramp rate and duration, specifying a power swing or ramp end over parameterizes the ramp event but are useful variables in analysis. This paper addresses the *detection* and *empirical modeling* of wind power ramps and their parameters.

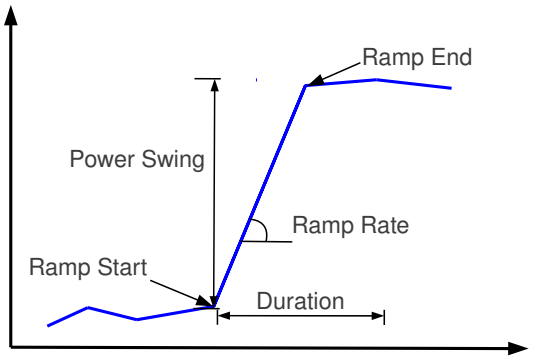

Fig. 1. Sample wind power signal. Indicated are items ramp detection algorithm aims to identify correctly. Ramp start and end times, as well the power change in interval as well as ramp rate. These variables are basis of statistical model.

The detection problem is formulated as follows. An input time series $X = \{(t_1, p_1), \ldots, (t_N, p_N)\}$ which represents a set of time and power pairs is given. Each interval $E = (i, j)$ represented indices $i, j : 1 \leq i < j \leq N$ into the the time series, as well as a possible ramp starting at i and ending at j. A ramp rule $R : E \times X \rightarrow \{0, 1\}$ maps intervals of X to a decision on whether the interval constitutes a ramp. The *optimal ramp detection* problem is to design an algorithm that recovers the sequence of intervals $E_1, \ldots E_k$ each corresponding to a ramp.

The empirical ramp modeling problem uses the detected sequence of intervals to identify and compute various ramp features useful for system analysis and forecasting. In the next subsections we develop the detection and modeling problems in more detail.

A. Optimal Ramp Detection

Given an input power time series $X = \{(t_1, p_1), \ldots, (t_N, p_N)\}$ and ramp rule $R : E \times X \rightarrow \{0, 1\}$, a ramp detection algorithm will output an interval set $\mathcal{E}_{X,R} = \{E_1, \ldots, E_L\}$ representing L detected ramps where each interval $E_i = (s_i, e_i)$ corresponds to the start (s_i) and end (e_i) of the i'th ramp event. Since the time series X is a constant throughout the paper, $\mathcal{E}_R$ is used interchangeably with $\mathcal{E}_{X,R}$. Consistently identifying ramps in data requires a clear notion of optimality of a set of candidate ramps. For any input X and ramp rule R an optimal set of intervals ${\mathcal{E}^\star}_{X,R}$ can be defined by:

Definition The optimal interval set for a ramp rule R is a set of intervals ${\mathcal{E}^\star}_R$ such that $\forall\,(s_l, e_l) \in {\mathcal{E}^\star}_R\ R(s_l, e_l) = 1$ however $\forall s', e'$ where $s' < s_l < e_l < e'$, $R(s', e_l) = 0$ and $R(s_l, e') = 0$.

If there existed such an s' or e', where $R'(s', e_l) = 1$ or $R'(s_l, e') = 1$, then an assignment $(s_l, e_l) \leftarrow (s', e_l)$ or $(s_l, e_l) \leftarrow (s_l, e')$ would update the intervals thus making them optimal. Also, we can say one interval set is equivalent to an-

other, $\mathcal{E}_{X,R'} = \mathcal{E}_{X,R}$ if $\forall\,(s'_l, e'_l) \in \mathcal{E}'_{X,R}$ and $\forall\,(s_l, e_l) \in \mathcal{E}_{X,R}$, $s'_l = s_l$ and $e'_l = e_l$.

1) Wind Ramp Rules: Implementing an optimal ramp detection method in practice requires specifying the rule sets R. We use $R\,(E)$ or $R\,(i,j)$ denote a rule applied to X in a signal interval $E = (i,j)$ in the remainder of the paper. In [8] and [6], three rules are proposed to determine whether an interval of the wind power time series is a ramp event:

$$R_0\,(i,j) = \mathbf{1}_{\{p_j - p_i > P_{val}\}} \quad \text{power swing thresholding} \tag{1}$$

$$R_1\,(i,j) = \mathbf{1}_{\{\max(p_i,\ldots,p_j) - \min(p_i,\ldots,p_j) > \alpha\}} \quad \text{max/min power difference thresholding} \tag{2}$$

$$R_2\,(i,j) = \mathbf{1}_{\left\{\frac{p_j - p_i}{t_j - t_i} > \alpha\right\}} \quad \text{ramp rate thresholding} \tag{3}$$

Eq. (1) is a power swing threshold test which checks whether the wind power has increased by a specified amount in a given time span. The parameter P_{val} specifies the threshold, which Δ species the time window for detecting such swings.

Eq. (3) checks whether the maximum and minimum observed power within an interval is greater than a threshold value α. It is interesting to note two features of this rule in relation to that in Eq. (1). First, Eq. (3) is more sensitive to fluctuations in the wind power signal and will have a high false positive rate in detecting ramps. Second, a signal satisfying Eq. (1) will always satisfy Eq. (3).

Eq. (3) checks whether the rate of increase for a specified interval is greater than a specified value. α is the power ramp rate threshold used.

The previous rules do not exclude intervals that contain sudden drops in power or large fluctuations inside an interval. Such fluctuations may determine a premature termination or start for a ramp. We use an additional rule that can be computed efficiently by tracking current interval maximum to overcome this issue:

$$R_c\,(i,j) = \prod_{m=i}^{j} \mathbf{1}_{\{p_m > \beta\ \max(p_i,\ldots,p_m)\}}, \quad \beta < 1. \tag{4}$$

Additional rules can be formulated. The generalized rule in Eq. (5) allows any definition to be considered.

$$R\,(i,j) = R_c\,(i,j) \prod_{r=1}^{M} R_r\,(i,j)\,. \tag{5}$$

Here $R_c\,(i,j)$ is given by Eq. (4) and $R_r\,(i,j)$ is given by any definition of a wind ramp.

B. Ramp Model and Characterization

Large datasets can be analyzed with an automatic ramp detection algorithm. Multiple sites and multiple years can be quickly processed to generate large numbers of ramp intervals. Based on these empirically detected ramps we need to properly characterize and study the statistics of ramp rates, durations, and start times. Currently, such analysis is mostly ad-hoc and visual. The statistics of ramps can also be used to create generative models of wind power time series, that can be later used in forecasting and simulation. A generative model of a stochastic process, is a Bayesian probabilistic description of a phenomenon used to randomly generate observations for simulation or testing. Models that incorporate ramp forecasting parameters can also be integrated into stochastic optimization problems for solving dispatch problems in wind power.

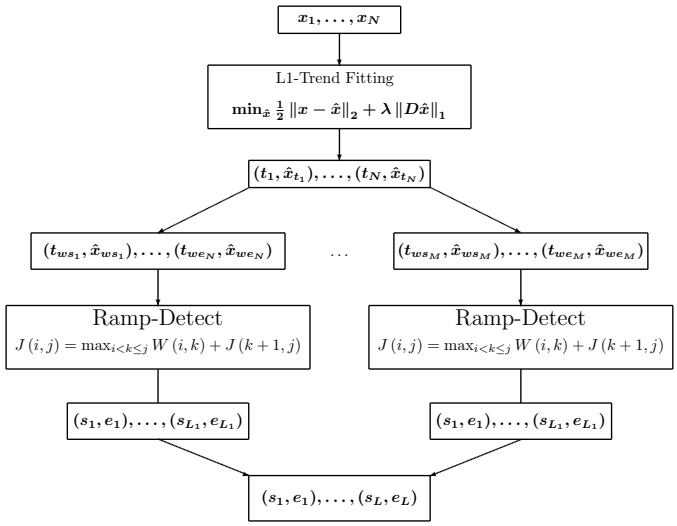

Fig. 2. Ramp detection methodology flowchart shows trend fitting, sliding window and dynamic programming recursion used.

IV. Optimal Ramp Detection

The definition of optimality in Section III-A motivates identifying ramps as a problem of detecting optimal intervals of the time series satisfying a given set of ramp rules. Section IV-A poses this problem as a dynamic program. To emphasize trends in the data at the appropriate timescales, a piecewise linear trend fitting technique is applied to the data prior to the dynamic program. This technique is presented in subsection IV-C. A sliding window approximate technique is proposed which makes the problem tractable for arbitrarily long inputs X and is described in subsection IV-D. The complete algorithm combining trend fitting, dynamic programming solution and sliding window approximation is shown in Figure 2 and is referred to as L1-Ramp Detect with Sliding Window.

A. Dynamic Program Formulation

The definition of optimal ramp intervals (Definition III-A) naturally implies that the longest sequence of points that satisfy the ramp rules is a ramp event. The starting point and end point of such a sequence forms an interval. Directly identifying this set of longest intervals would require evaluating the ramp rules on each subsequence of the original dataset. Instead, this problem can be formulated as a dynamic program by realizing that every subinterval of an interval that satisfies ramp rules will also satisfy ramp rules. If every interval that satisfies ramp rules is assigned a *score* that is an increasing function of the interval length, then the problem reduces to finding a set of intervals that maximizes this score.

Given an appropriate scoring function $W : E \times X \to \mathbb{R}^+$ on intervals of the time series, the optimal ramp start and stop

times are recovered by maximizing the objective function J according to the dynamic program:

$$J(i,j) = \max_{i<k\leq j} W(i,k) + J(k+1,j) \tag{6}$$

$J(i,j)$ is the maximum attained score in signal interval (i,j). It can be computed as the maximum over $j-i$ subproblems. Maximizing over all subproblems yields an optimal overall solution since each subproblem is in terms of a maximum score in the interval $(k+1,j)$, and $W(i,k)$ is the (positive) weight given to the interval (i,k) .

The solution to Eq. (6) can be represented as $J^{\star}(i,j) = \sum_{l=1}^{L^*} W(s_l^*, e_l^*)$, where L^* is the optimal number of subsequences or ramp events. s_l^*, e_l^* are the start and end indices of the l'th subsequence. Therefore, we can define for every score function an interval set $\mathcal{E}^{\star}{}_W = \{(s_1^{\star}, e_1^{\star}) \ldots, (s_{L^{\star}}^{\star}, e_{L^{\star}}^{\star})\}$ which solves the recursion in Eq. (6). Algorithm 1 is a simple standard recursion to solve Eq. 6.

B. Ramp Score Function

The dynamic program defined earlier requires a *proper* cost function that evaluates the cost of each subsequence. For the interval set $\mathcal{E}^{\star}{}_W$ solving Eq. (6) to be equal to the optimal ramp sequence assignment $\mathcal{E}_R$ for rule R, the score function should satisfy a super additivity property (see Appendix A):

$$\begin{aligned} \forall i<k<j : R(i,j) &= 1, \\ W(i,j) &> W(i,k) + W(k+1,j) \end{aligned} \tag{7}$$

Since the optimal solution will extract indices where $W(s_l, e_l)$ has high positive score, we desire the indices to be where ramps start and end. Therefore we want to score ramp events and penalize non ramp intervals. Also, in order to prevent arbitrary subintervals of a ramp interval to receive more weight than the entire interval, the super-additivity condition is imposed. The condition defines an entire family of weight functions. One choice is shown in Eq. (8). Any other choice satisfying Eq. (7) could be used. All choices recover the same set of ramp intervals.

$$W(i,j) = (j-i)^2 \mathbf{1}_{\{R(i,j)=1\}} \tag{8}$$

C. Trend Fitting

The wind power time series contains uncertainty and underlying structure at multiple time scales. Investigating phenomenon at appropriate time scales requires elimination of noise or trends outside the time scale of interest. For the case of wind ramp events, the time scale of interest is on the order of minutes to hours. Additionally, the algorithm presented in Section IV-A requires $O\left(N^3\right)$ computations and $O\left(N^2\right)$ memory size which is prohibitive for large multiple year and high resolution time-series. It is important then reduce data size while still preserving appropriate trends in the data set.

Trend filtering (or fitting) addresses this issue and is capable of removing short term fluctuations in wind power, while exposing trends in data. The proposed ramp detector uses a piecewise linear trend fitting as a preprocessing step. A simple and tractable technique for generating piecewise linear fits to data is obtained by solving the convex program [10]:

$$\min_{\hat{x}} \frac{1}{2} \|x - \hat{x}\|_2 + \lambda \|D\hat{x}\|_1 , \tag{9}$$

where we assume $x = \{p_1, \ldots, p_N\}$ and $t_i = i$ and the transformation D is the second derivative operator. The $\|D\hat{x}\|_1$ enforces sparsity in the second derivative, and subsequent piecewise linearity in the approximation $\hat{x}$. Sifting through multiple time scales is done by changing the parameter λ. The choice of λ is subjective and can affect the detected ramp events. After performing the optimization, the set of piecewise lines are extracted from $\hat{x}$ by thresholding the second derivative by some value γ. Since this is an approximate algorithm, care should be taken in choosing a threshold value. The resulting thresholded piecewise signal is given by:

$$\hat{X}_{PW} = \{(t_1, \hat{p}_{t_1}), \ldots, (t_T, \hat{p}_{t_T})\}, \tag{10}$$

where $(t_1, \hat{p}_{t_1}), \ldots, (t_T, \hat{p}_{t_T})$ represent points which satisfy $\|D\hat{x}\|_1 > \gamma$.

Notice that trend filtering via Eq. (9) and detecting drops in the wind power via Eq. (4) deal with fluctuations in power on different time scales. Trend filtering removes very short term fluctuations, but Eq. (4) will penalize drops in power in a longer timescale.

Algorithm 1 Optimal Ramp Detection Algorithm.

```
L ← length(p) {Initialize Score of zero length segments.}

for i = 1 → L do
    J[i,i] ← 0
end for
// Compute optimal cost for segments.
for n = 2 →L do
    for i = 1 → L − n + 1 do
        j ← i + n − 1
        for k = i → j − 1 do
            q ← W(i,k) + J(k,j)
            if q > J[i,j] then
                J[i,j] ← q
                K[i,j] ← k
            end if
        end for
    end for
end for
```

D. Sliding Window

Further reduction in total computation time is obtained by processing overlapping sections (windows) of the time series separately or in parallel and combining the detected ramps in a post processing step. The optimality of the original procedure can be maintained by adding a maximum duration parameter for a ramp. The distributed procedure is capable of detecting all ramps if each window lasts longer than the longest ramp duration.

The approximate piecewise signal Eq. (10) is split into a number of overlapping windows, parameterized by window

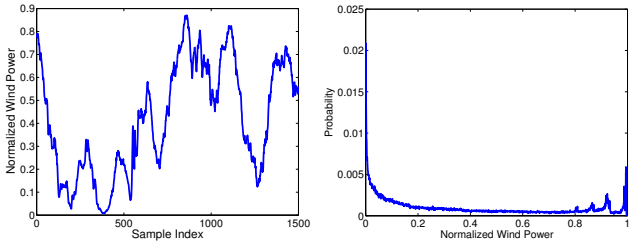

Fig. 3. **(Left)** Sample interval of wind power time series from Bonneville Power Authority. **(Right)** Empirical histogram of total power production over entire year.

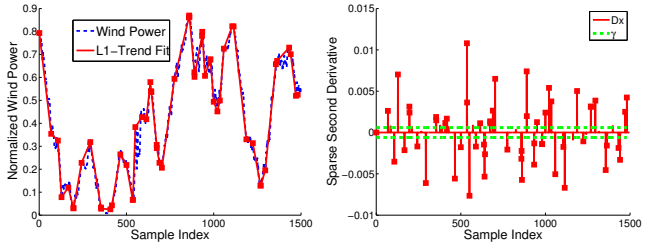

Fig. 4. **(Left)** Wind time series and trend fit signal. **(Right)** L1 Trending leads to sparse second derivative from L1 constraint on Dx enforcing sparsity.

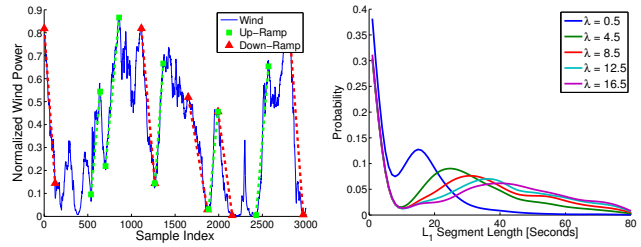

Fig. 5. **(Left)** Wind power timeseries with detected wind ramps. **(Right)** Empirical distribution for l_1 trend fitting segments. Choice of λ used as tunning parameter to uncover trends in data at appropriate timescales.

length, WL and window overlap WO. The result of this splitting is $M = \lfloor \frac{N+WL-WO}{WL-WO} \rfloor - 1$; subsets of X_{pw} of the form $X_{ws_1:we_1}, \ldots, X_{ws_M:we_M}$ where window start and end indices are given by Eq. (12).

$$ws_i = (i-1)(WL - WO) \tag{11}$$

$$we_i = (i)\,WL - (i-1)\,WO \tag{12}$$

Each application of the ramp detection algorithm will yield the set of detected ramp events for each windowed signal $\mathcal{E}_{X_{ws_1:we_1},R}$, $\ldots$, $\mathcal{E}_{X_{ws_M:we_M},R}$. These results are merged, ensuring that ramps in the regions where windows overlap are uniquely determined by choosing the longest ramp, to yield the set of detected ramp events ($\mathcal{E}_{X,R}$).

V. Experimental Results

A. Data Description

The optimal ramp detection scheme and subsequent statistical modeling are applied to two different sets of data. The first being a time series of actual wind power production from the Bonneville Power Authority (BPA). The second is a set of $2,882$ simulated wind power production sites from the NREL Western Integration Dataset. The first data set contains $15,768,000$ samples representing measured wind turbine power output sampled every 2 seconds. The data represents power generation from January 1st, 2005 to December 31st, 2006 at a selected facility. The capacity of the plant is 24 MW. The second set is the Western Integration Study [13] dataset that contains 30,000 time series sampled every 10 minutes. It covers the entire western half of the United States, however we are only interested in the California production. Each time series represents the potential wind power generated from 10 Vestas V-90 generators with a rated capacity of 3 MW. Thus at each site, the maximum production is 30 MW. The data used contains only the wind generation sites from the state of California.

B. Data Preprocessing

The BPA dataset requires extensive preprocessing since it is data from an actual site. First the data sets are normalized to the nameplate capacity of the system. Then, negative recorded power values are replaced with zero power output, since they occur when the power output is low. Then, the data is interpolated to correct for uneven sampling timestamps of the original data. Finally, the data is downsampled by 100 resulting in a effective sample rate of 200 seconds. The NREL dataset only required normalizing each series to the nameplate capacity.

C. Piecewise Linear Trend Fitting

As described in Section IV-C, a trend filtering algorithm is applied to the normalized power time series, by optimizing a penalized least squares with penalty parameter λ. Low values of λ produces a good fit with low approximation error at the expense of a larger number of line segments (Figure 5). A large value of λ produces a coarse fit, but results in very few line segments. For the remainder of this work, a piecewise signal was extracted with a λ value of 0.5. For $\lambda > 0.5$, the piecewise fit tends to deviate from the overall trend of the signal. The value was chosen by visual inspection of the piecewise signal. The thresholding parameter of the second derivative, $\gamma = 0.0001$ was also chosen by visual inspection. The piecewise fit and the sparse second derivative of the BPA dataset are shown in Figure 4. The approximate signal has a size $|\hat{X}_{PW}| = 13,217$ segments (as compared to $15,768,000$ samples).

D. Ramp Detection

A ramp detection experiment was performed with the ramp rules outlined in Section III-A1. The rules in Eqs. (1),(3) and (4) have tuning parameters set to $P_{val} = 0.25$, $\alpha = 0.003$, $\beta = 0.9$. The P_{val} corresponds to $6.0MW$ of increased power for detected ramps while α corresponds to 0.004 MW/hour minimum ramp rate. The windowed dynamic programming methodology as described in Sections IV and IV-D is used for ramp detection with a window length $WL = 2000$ and window overlap $WO = 500$. For the BPA dataset, the input vector of length 13217 points representing a year of data was

	Manual Detected	Manual Not Detected
Alg. Detected	42	8
Alg. Not Detected	5	

TABLE I
ACCURACY OF RAMP DETECTION BY COMPARING MANUALLY AND ALGORITHM DETECTED RAMPS.

used. The computation was completed in 3.2 seconds on a 1.67 GHz dual core CPU with 2G of RAM. The software was implemented in the C programming language. For a year long dataset, a total of 546 up-ramps were detected as well as 552 down-ramps. A sample of the original power time series and the detected up and down ramps is shown in Figure 5. For the NREL dataset, ramp detection was performed for each of the $2,882$ sites and similar performance was observed. We should note, quite remarkably that about to 25% of the BPA timeseries is classified to be in either ramping state.

E. Evaluating Accuracy of Ramp Detection

A simple test is performed to compare a manually selected set of ramps to those identified by our technique on the BPA dataset. The first $13,000$ samples of the original wind time series were visually inspected to mark all regions of large positive or negative deviation. A total of 50 up and down ramps were identified. An attempt was made to indicate large swings which were abrupt and had large change of power. In contrast, a total of 47 ramps were detected by the algorithm, using the parameters defined in Section V-D. The set of ramps detected by the computer and by visual inspection are denoted by C and H respectively. Ramps in both sets are compared to identify equivalent ramps. Two ramps are considered equivalent, if the overlapping intervals between the two is greater than 80% of the mean length of the two intervals. An accuracy matrix comparing manually detected and algorithm detected ramps is computed based on this criteria by identifying ramps detected by both methods, and by each method alone (Table V-E).

Ramps detected by manual inspection but missed by the algorithm are of interest, since they are potential missed ramp events. There were 5 such ramps. A closer inspection of these missed events revealed a mismatch between the subjective rules used by the human to detect ramps, and the parameters used by the algorithm to detect ramps. In these situations, intervals with deviations slightly less than the threshold value P_{val} were counted as ramps while our algorithm discarded these intervals. Adjusting the thresholds could recover these ramps, but this disagreement also reflects the lack of consensus the meaning of a ramp rule.

VI. Empirical Statistics of Wind Ramps

In this section, we present various statistics to analyze the detected ramp events and parameterize the ramp event process on the BPA and NREL datasets. Such statistics can be computed for other data sources to reveal various important properties of the ramping process. We initially focus on the BPA data to illustrate various metrics, and then evaluate the NREL data set.

A. Distribution of Ramp Rate, Duration and Swing

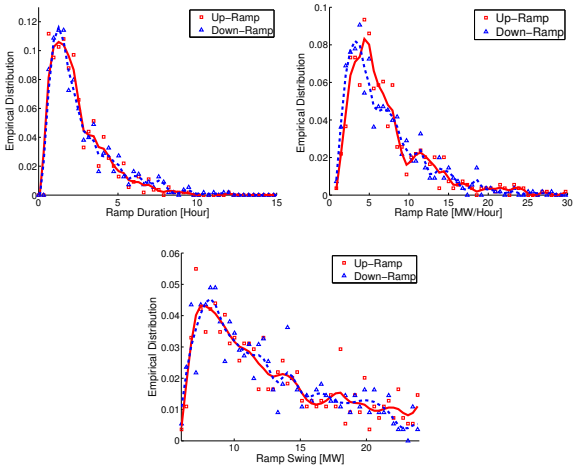

Fig. 6. Empirical distributions for ramp duration, rate and swing. Note matching of distributions for up-ramps and down-ramps.

We compute empirical probability density functions (pdf) for ramp duration, ramp rate and swing for up and down ramps. The pdf is computed by applying a smoothing spline to histograms of the data. Results are shown in Figure 6 and show a high degree of symmetry between the up ramp and down ramp parameters.

A mean ramp duration of 2.5 hours and minimum duration of 30 minutes were observed for both ramp types. 95% of ramps were less than 5.6 hours for both cases. Note the heavy tailed nature of the data: most ramp durations are centered near the mean value with longer ramp events occurring infrequently. For both up and down ramp types, a mean ramp rate of 1.2 MW/hour was observed. The 50% interquantile range for both ramp rates is 0.7 to 1.5 MW/hour and 95% of rates are less than 2.4 MW/hour. The distribution also seems fat tailed. For ramp swing, a mean swing of around 12.2 MW was observed with a minimum of 6 MW. The minimum observed swing corresponds to the maximum power output multiplied by the minimum swing threshold P_{val}. This is an expected consequence of the ramp detection algorithm. A maximum swing of 24 MW is also observed which corresponds to the nameplate capacity of the farm. The 50% interquantile range of up ramp swings is 15 to 20 MW while 95% of swings are less than 21 kW.

B. Joint Distributions of Ramp Duration with Ramp Rate and Swing

Joint distributions for ramp duration and rate as well as ramp duration and swing are shown in Figure 7. The joint distribution of rate and duration shows a clear inverse relationship.

This is expected for the following reason. Ramp rate defined as $\frac{swing}{duration}$ is plotted against $duration$. Since the variable $swing$ can be bounded: $P_{val} < swing < 41KW$, we can bound the ramp rate as well. More precisely, $\frac{P_{val}}{duration} < rate\,(duration) < \frac{24MW}{duration}$.

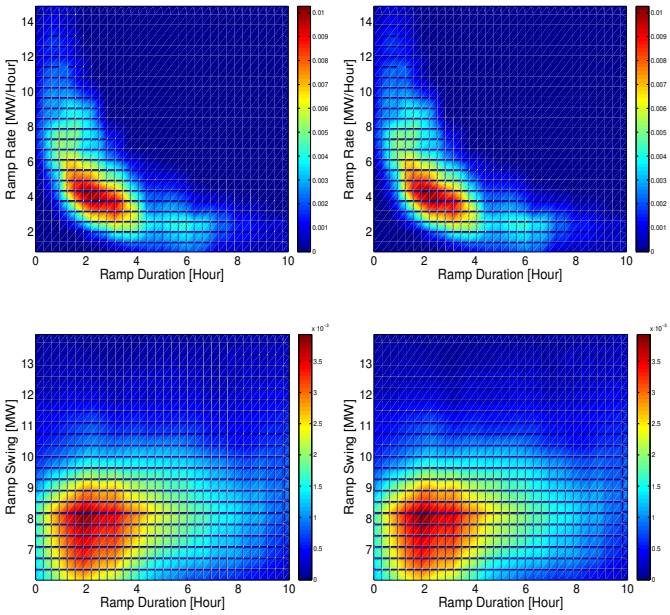

Fig. 7. **(Top)** Joint distribution of ramp duration and ramp rate for up and down ramps. **(Top)** Joint distribution of ramp duration and ramp swing for up and down ramps.

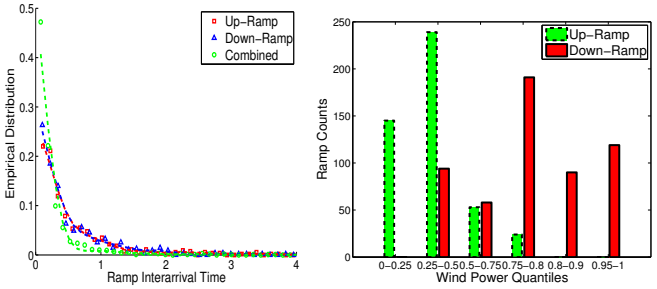

Fig. 8. **(Left)** Interarrival time distributions for up ramp, down ramp and combined time sequences. **(Right)** Ramp counts conditioning on wind power quantiles.

We see that there is less dependence between ramp swing and duration as opposed to duration and rate. We also see a distinct mode of 12 MW of ramp swing and 2 Hours of ramp duration. These correspond to the modes of duration and swing shown in Figure 6

1) Interarrival Time: Ramp start times indicate the time a ramp 'arrives' at a site. When a sequence of ramps arrives, the process of arrival can be characterized by the time difference in arrivals or interarrival times. Distributions of such arrival times should be captured by any wind model and can give important information for operations planning, for example, to assess ramping requirements for generators. Up ramps are always followed by down ramps. We define up (down) ramp interrarrival times as the arrival time difference between two consecutive up (down) ramps, and a combined interarrival time as the interarrival time between an up ramp and the subsequent down ramp. Figure 8 displays interarrival time distributions.

The mean interarrival time observed was 0.33 days or about 8 hours. The 50% interquantile range of combined ramp interarrival times are between 1.8 and 7.2 hours in length while 95% of interarrival times are less than 27 hours long. The minimum observed interarrival time was 30 minutes in length. This an important property in potential modeling applications. A typical statistical model for arrival processes is a Poisson process, that assumes an exponential interarrival time distribution, which includes infinitesimally small interarrival times. Clearly, this is not the case. This is a natural consequence of the ramp detection algorithm. If two up ramp events are very close together, the dynamic program will combine them and report a single larger wind power ramp.

C. Seasonality of Ramp Counts

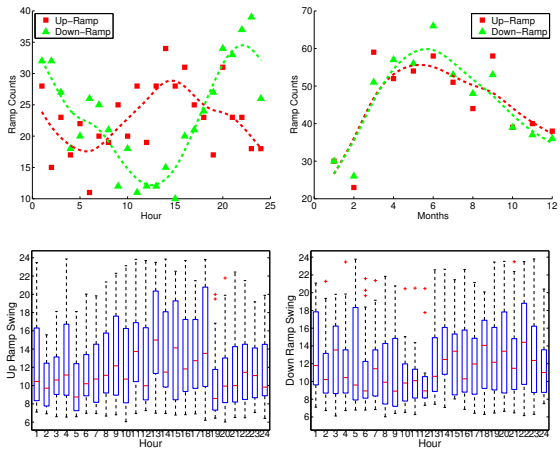

Fig. 9. (Top) Seasonality of ramp events. This is in the form of daily cycles corresponding to short term weather patters as well as yearly cycles corresponding to changes in the seasons. (Bottom) Hourly distributions of Ramp Swings.

Figure 9 illustrates the seasonality in the detected ramps. Ramps are frequent and average 1 to 2 ramps a day. We see first that ramp events occur much more often in the spring and summer than in the fall and winter. In fact, they increase by a factor of two. Second, we notice that up ramp and down ramp events tend to occurs at different parts of the day. Up ramps tend to occur more often during the mid-day and afternoon periods, while down ramps occurs during the evening and night times. These are consequences of two different weather cycles. The first being the changes in the seasons and the second is due to daily weather patterns. Hourly up and down ramp swing distributions are shown in Figure 9. The figures show a dependence of the distribution of ramp swing on hour of the day. The time dependent mean's and variances are useful for computing the reserve capacity of the power grid.

D. NREL Western Wind Data Set

1) Aggregate Statistics: For the 2882 locations of simulated wind power generation we run the ramp detection algorithm and report the summary distributions of ramp swing, duration and ramp. A total of 674650 ramps were detected and used to form the following empirical distributions. Figure 10 shows the aggregated statistics of each ramp event from all 2882

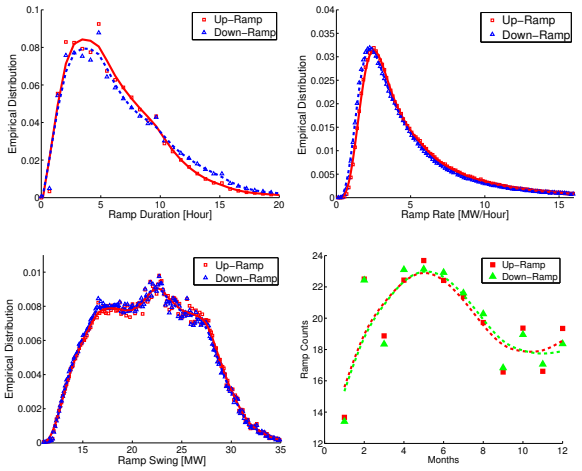

Fig. 10. Empirical distributions for ramp duration and rate for up-ramps and down-ramps. Also monthly ramp counts.

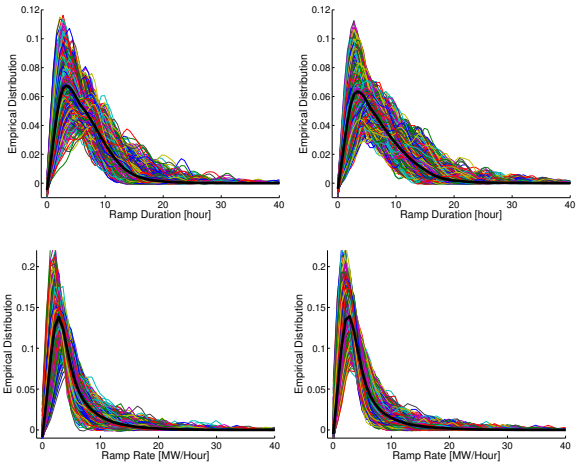

Fig. 11. Empirical distributions for ramp duration for up-ramps **(Left)** and down-ramps **(Right)**. The superimposed black line shows the mean histogram of averaging over every site.

locations over the course of a year. The mean duration is 5.5 hours which significantly larger than the value determined in the BPA dataset. The aggregate duration distribution has a fatter tail than that of the BPA data. Also, the ramp swing distribution has a symmetric shape centered around 24 MW which is very different than the BPA single site distribution. This discrepancy illustrates the difference between true power generation data versus that of simulation. The ramp swing in the observed dataset again shows a fat tailed nature, while the simulation data does not. Figure 11 shows the individual empirical distributions of the ramp duration and rate for all 2882 sites. These common shapes, shows that we can model individual sites by not only their generation timeseries but by a reduced number of parameters representing families of distributions for rate, swing and duration. Figure 10 illustrates monthly seasonality in ramp counts. The plot shows the mean ramps detected in per generation site. We see a similar increase of ramp counts in the summer months much like in the BPA dataset. However, the total number of counts is much less for the NREL dataset. While a month produces up to 70 ramp events in the BPA dataset, the California average is much less. Also, the variation between winter months (where observed ramp counts are lowest, and summer months (where they are highest) is much less. In the BPA data, there was a two fold difference between the two, while the CA data shows close to a 75% increase between the winter and summer ramp counts.

VII. Conclusion

We have described an optimal method for detecting wind ramp events for varying lengths under any arbitrary rule set. A dynamic programming recursion with L_1 trend fitting is used to optimally segment wind power data to ramp events and non-ramp events. The technique is shown to be optimal, and is used to provide empirical statistics of wind ramps for real data. The data is then used to propose a simple model for characterizing ramp events as well as outline future directions for ramp event based short term forecasting.

References

[1] C. Barber, J. Bockhorst, and P. Roebber. Auto-regressive hmm inference with incomplete data for short-horizon wind forecasting. In J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R.S. Zemel, and A. Culotta, editors, *Advances in Neural Information Processing Systems 23*, pages 136–144. MIT Press, 2010.

[2] E. Y. Bitar, R. Rajagopal, P. P. Khargonekar, K. Poolla, and P. Varaiya. Bringing wind energy to market. *Power Systems, IEEE Transactions on*, PP(99):1, 2012.

[3] A. Bossavy, R. Girard, and G. Kariniotakis. Forecasting uncertainty related to ramps of wind power production. In *Sci. Proc. of the Eur. Wind Energy Conf.*, pages 20–23, 2010.

[4] J. DeCesaro, K. Porter, and M. Milligan. Wind energy and power system operations: A review of wind integration studies to date. *The Electricity Journal*, 22(10):34–43, 2009.

[5] R. Doherty and M. O'Malley. A new approach to quantify reserve demand in systems with significant installed wind capacity. *Power Systems, IEEE Transactions on*, 20(2):587–595, 2005.

[6] C. Ferreira, J. Gama, L. Matias, A. Botterud, and J. Wang. A survey on wind power ramp forecasting. Technical report, Argonne National Laboratory (ANL), 2011.

[7] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl. The state-of-the-art in short-term prediction of wind power: A literature overview. Technical report, ANEMOS. plus, 2011.

[8] C. Kamath. Understanding wind ramp events through analysis of historical data. In *Transmission and Distribution Conference and Exposition, 2010 IEEE PES*, pages 1–6. IEEE, 2010.

[9] C. Kamath. Associating weather conditions with ramp events in wind power generation. In *Power Systems Conference and Exposition (PSCE), 2011 IEEE/PES*, pages 1 –8, march 2011.

[10] S.J. Kim, S. Boyd, and D. Gorinevsky. L-1 trend filtering. *SIAM review*, 51(2):339–360, 2009.

[11] M. Lei, L. Shiyan, J. Chuanwen, L. Hongling, and Z. Yan. A review on the forecasting of wind speed and generated power. *Renewable and Sustainable Energy Reviews*, 13(4):915–920, 2009.

[12] JC. Palomares-Salas, JJG. De la Rosa, JG. Ramiro, J. Melgar, A. Agera, and A. Moreno. Comparison of models for wind speed forecasting. In *Proc. of Int. Conf. on Comp. Sci. (ICCS 2009)*, pages 25–27, 2009.

[13] C.W. Potter, E. Grimit, and B. Nijssen. Potential benefits of a dedicated probabilistic rapid ramp event forecast tool. In *Power Systems Conference and Exposition, 2009. PSCE'09. IEEE/PES*, pages 1–5. IEEE, 2009.

[14] P. Varaiya R. Rajagopal, E. Bitar and F. Wu. Risk limiting dispatch: A framework for dispatching random power. *IEEE Transactions on Automatic Control*, 2011. "submitted".

[15] P. Varaiya, F. Wu, and J.W. Bialek. Smart operation of smart grid: Risk-limiting dispatch. *Proceedings of the IEEE*, 99(1):40–57, 2011.

[16] H. Zareipour, H. Dongliang, and W. Rosehart. Wind power ramp events classification and forecasting: A data mining approach. In *Power and Energy Society General Meeting, 2011 IEEE*, pages 1 –3, july 2011.

[17] H. Zheng and A. Kusiak. Prediction of wind farm power ramp rates: A data-mining approach. *Journal of Solar Energy Engineering*, 131:031011–1, 2009.

APPENDIX

Non-event interval set $\overline{\mathcal{E}} = \left\{\overline{E_1}\ldots,\overline{E_L}\right\}$ is the set of intervals $\overline{E_l} = (\overline{s_l},\overline{e_l})$ constructed from event interval set $\mathcal{E}$ with $\forall\,(s_l,e_l)\,,(s_{l+1},e_{l+1}) \in \mathcal{E} : 1 \leq s_l, e_{l+1} < N$ assigning $(\overline{s_l},\overline{e_l}) \leftarrow (e_l+1, \overline{s_{l+1}}-1)$.

Note that we can represent with any rule R an event sequence $\mathcal{S}_R = \left\{\ldots,\overline{E_l},E_l,\overline{E_{l+1}},E_{l+1},\ldots\right\}$ which is the optimal arrangement of ramp events and non-ramp events for a given input signal X.

We state the following properties without proof.

(P1) $\forall\,(\overline{s_l},\overline{e_l}) \in \overline{\mathcal{E}}$ and $\forall i,j : \overline{s_l} \leq i < j < \overline{e_l}$ $R\,(i,j) = 0$ and likewise from 7 $W\,(i,j) = 0$

(P2) For any three points, $x_i \notin X_{E_l}$, x_j, $x_k \in E_l$ with the following must hold: if $R'\,(x_i,\ldots,x_j) = 0$ and $R'\,(x_j,\ldots,x_k) = 1$ then $R'\,(x_i,\ldots,x_k) = 0$

(P3) $\forall i,k,j : 1 < i < k < j < N$, $J^\star(i,j) \geq J^\star(k,j)$

Theorem A.1: If W satisfies the super additivity property in Eq. (7) and input X to Eq. (6) has an event sequence $\mathcal{S}_R = \left\{E_1,\overline{E_1},\ldots,E_L,\overline{E_L}\right\}$, the solution to Eq. (6) is $J^\star\,(1,N) = \sum_{l=1}^{L} W\,(s_l,\ e_l)$ where $(s_l,\ e_l) \in \mathcal{E}$.

Proof: The proof will proceed inductively with the following.

(i) If signal $x_{s_l},\ldots x_{e_L}$ with event sequence $\mathcal{S}_R = \left\{E_l,\overline{E_l},\ldots,E_L,\overline{E_L}\right\}$ has as solution to Eq. (6) being $J^\star\,(s_l,e_L) = \sum_{i=l}^{L} W\,(s_i,\ e_i)$, then the solution to $x_{\overline{s_{l-1}}},\ldots,x_{\overline{e_l}},x_{s_l},\ldots,x_{e_L}$ with event sequence $\mathcal{S}_R = \left\{\overline{E_{l-1}}E_l,\overline{E_l},\ldots,E_L,\overline{E_L}\right\}$ is again $\sum_{l=l}^{L} W\,(s_l,\ e_l)$.

(ii) If signal $x_{\overline{s_{l-1}}},\ldots,x_{e_L}$ with event sequence $\mathcal{S}_R = \left\{\overline{E_{l-1}},E_l,\overline{E_l},\ldots,E_L,\overline{E_L}\right\}$ has as solution to Eq. (6) $J^\star\,(\overline{s_l},e_L) = \sum_{l=l}^{L} W\,(s_l,\ e_l)$, then the solution to $x_{s_{l-1}},\ldots,x_{e_{l-1}},x_{\overline{s_{l-1}}},\ldots,x_{e_L}$ with event sequence $\mathcal{S}_R = \left\{E_{l-1},\overline{E_{l-1}},E_l,\overline{E_l},\ldots,E_L,\overline{E_L}\right\}$ is $W\,(s_{l-1},\ e_{l-1}) + \sum_{l=l}^{L} W\,(s_l,\ e_l)$.

(iii) **Base Case** The solution to sequence $x_{s_L},\ldots,x_{e_L}$ with event sequence $ES = \{E_L\}$ is $W\,(s_L,\ e_L)$

(i) First show that $\forall i : s_{l-1} < i < e_{l-1}$ $k_1^* < s_l$ where $k_1^* = \arg\max_{i<k<e_L} W\,(i,k) + J\,(k+1,e_L)$. Assume $k_1^* \geq s_l$, then from (P1) $W\,(i,k) = 0$ and from (P3) $J\,(k+1,e_L) < J\,(s_l,x_{e_L})$ leading to contradiction. We can unfold the recursion as follows.

$$J^\star\,(\overline{s_{l-1}},e_L) = \max_{\overline{s_{l-1}}<k_1<s_l} W\,(\overline{s_{l-1}},k_1) + J\,(k_1+1,e_L) \tag{13}$$

$$= \max_{\overline{s}_{l-1}<k_1<s_l} W\,(\overline{s_{l-1}},k_1) + \tag{14}$$

$$\max_{k_1<k_2<s_l} W\,(k_1,k_2) + J\,(k_2+1,e_L) \tag{15}$$

$$= \max_{\overline{s_{l-1}}<k_1<s_l} W\,(\overline{s_{l-1}},k_1) + \max_{k_1<k_2<s_l} W\,(k_1,k_2) \tag{16}$$

$$+\ldots+ \max_{k_i<k_{i+1}<s_l} W\,(k_i,k_{i+1}) + J\,(k_i+1,e_L) \tag{17}$$

$$= \max_{\overline{s_{l-1}}<k_1<k_2<\ldots<k_{i-1}<k_i<s_l} W\,(\overline{s_{l-1}},k_1) + \tag{18}$$

$$W\,(k_1,k_2)+\ldots+W\,(k_{i-1},k_i) + J\,(k_i+1,e_L) \tag{19}$$

$$= \max_{\overline{s_{l-1}}<k_1<k_2<\ldots<k_{i-1}<k_i<s_l} J\,(k_i+1,e_L) \tag{20}$$

$$= J^\star\,(s_l,\ldots,e_L) \tag{21}$$

Eq. (13)-(19) is a consequence of evaluating the recursion for i steps and using the fact that $k_i < s_l$. From (P1) we can reduce Eq. (20) to (21). Finally we observe $J\,(\overline{s_{l-1}},e_L) = J\,(\overline{s_{l-1}}+1,e_L) = \ldots = J\,(k_i+1,e_L) = \ldots = J\,(\overline{s_l},e_L)$

(ii) First show that $\forall i : s_{l-1} < i < \overline{s_l}$ $k_1^* < \overline{s_l}$ where $k_1^* = \arg\max_{i<k<e_L} W\,(i,k) + J\,(k+1,e_L)$. Assume $k_1^* > s_l$, then from (P1) $W\,(i,k) = 0$ and from (P3) $J\,(k+1,e_L) < J\,(s_l,x_{e_L})$ leading to contradiction.

$$J^\star\,(s_{l-1},e_L) = \ldots \tag{22}$$

$$= \max_{s_{l-1}<k_1<\overline{s_l}} W\,(s_{l-1},k_1) + J\,(k_1+1,e_L) \tag{23}$$

$$= \max_{s_{l-1}<k_1<\overline{s_l}} W\,(s_{l-1},k_1) + \tag{24}$$

$$\max_{k_1<k_2<\overline{s_l}} W\,(k_1,k_2) + J\,(k_2+1,e_L) \tag{25}$$

$$= \max_{s_{l-1}<k_1<\overline{s_l}} W\,(s_{l-1},k_1) + \max_{k_1<k_2<\overline{s_l}} W\,(k_1,k_2) \tag{26}$$

$$+\ldots+ \max_{k_{i-1}<k_i<\overline{s_l}} W\,(k_{i-1},k_i) + J\,(k_i+1,e_L) \tag{27}$$

$$= \max_{s_{l-1}<k_1<k_2<\ldots<k_{i-1}<k_i<\overline{s_l}} W\,(s_{l-1},k_1) \tag{28}$$

$$+ W\,(k_1,k_2)+\ldots+W\,(k_{i-1},k_i) + J\,(k_i+1,e_L) \tag{29}$$

$$= W\,(s_{1-1},e_{l-1}) + \sum_{l=l}^{L} W\,(s_l,\ e_l) \tag{30}$$

Eq. (32)-(38) is a consequence of evaluating the recursion for i steps and using the fact that $k_i < \overline{s_l}$. And from the superadditivity property of W we have Eq. (39). **(iii)**

$$J^\star\,(s_L,e_L) = \ldots \tag{31}$$

$$= \max_{s_L<k_1<e_L} W\,(s_L,k_1) + J\,(k_1+1,e_L) \tag{32}$$

$$= \max_{\overline{s_L}<k_1<e_L} W\,(s_L,k_1) + \max_{k_1<k_2<e_L} W\,(k_1,k_2) \tag{33}$$

$$+ J\,(k_2+1,e_L) \tag{34}$$

$$= \max_{\overline{s_L}<k_1<e_L} W\,(s_L,k_1) + \max_{k_1<k_2<e_L} W\,(k_1,k_2) \tag{35}$$

$$+\ldots+ \max_{k_i<k_{i+1}<e_L} W\,(k_i,k_{i+1}) + J\,(k_i+1,e_L) \tag{36}$$

$$= \max_{\overline{s_L}<k_1<k_2<\ldots<k_{i-1}<k_i<e_L} W\,(s_L,k_1) + W\,(k_1,k_2) \tag{37}$$

$$+\ldots+W\,(k_{i-1},k_i) + J\,(k_i+1,e_L) \tag{38}$$

$$= W\,(s_L,e_L) \tag{39}$$

■