Optimal Maintenance Strategies for Wind Turbine Systems Under Stochastic Weather Conditions

Eunshin Byon, Student Member, IEEE, Lewis Ntaimo, and Yu Ding, Member, IEEE

Abstract

We examine optimal repair strategies for wind turbines operated under stochastic weather conditions. In-situ sensors installed at wind turbines produce useful information about the physical conditions of the system, allowing wind farm operators to make informed decisions. Based on the information from sensors, our research objective is to derive an optimal observation and preventive maintenance policy that minimizes the expected average cost over an infinite horizon. Specifically, we formulate the problem as a partially observed Markov decision process. Several critical factors, such as weather conditions, lengthy lead times and production losses, which are unique to wind farm operations, are considered. We derive a set of closed-form expressions for the optimal policy and show that it belongs to the class of monotonic four-region policies. Under special conditions, the optimal policy also belongs to the class of three-region policies. These structural results of the optimal policy reflect the practical implications of the turbine deterioration process.

Index Terms

Dynamic programming, partially observed Markov decision process, random deterioration, stochastic environments, wind turbine operations and maintenance.
SUMMARY AND CONCLUSION

SUMMARY - We address the problem of finding the optimal repair strategies for wind turbine systems using partially observed Markov decision process. A new maintenance model is presented, considering critical factors that are unique to wind turbine operations. Based on the analysis of the structural properties of the wind turbine maintenance model, we also devise efficient algorithms to help numerically attain the optimal policies.

CONCLUSION - Wind turbines are located in remote areas and operated under harsh, non-stationary conditions. Observations of wind turbine conditions are costly and uncertain. Lead time for repairing are typically longer than other machineries traditionally considered, and suitability of executing or continuing a repair job depends on stochastic weather condition. Revenue losses upon failures are significant. There has not been a model considering all these factors in the maintenance research of wind power systems; to the best of our knowledge, the proposed model is the first for wind turbine maintenance. We also provide analysis that generates the insights of the model structure and enables efficient numerical solutions. This work is of interest to the reliability community because we believe that using the proposed model leads to the practical operations and maintenance guidelines that reduces the repair costs and increases the marketability of wind energy generation.
ACRONYMS

NA no action
PM preventive maintenance
OB observation
CM corrective maintenance
O&M operations and maintenance
CBM condition-based monitoring
POMDP partially observed Markov decision process
AM4R At-Most-Four-Region
AM3R At-Most-Three-Region

NOTATION

$\pi$ information state
$\tilde{\pi}$ information state when system is stopped for repairs
$\pi'(\pi)$ information state at next decision period
$P$ transition matrix
$R(\pi)$ reliability
$e_i$ unit vector with 1 in $i$th position
$C_{CM}$ cost for CM
$C_{PM}$ cost for PM
$C_{OB}$ cost for OB
$T$ lead time
$\tau$ revenue losses per period
$W_{PM}, W_{CM}$ probabilities that adverse weather conditions occur
$V_n(\pi)$ minimum expected total cost-to-go
$NA_n(\pi)$ expected total cost-to-go when NA is taken
$PM_n(\pi)$ expected total cost-to-go when PM is taken
$OB_n(\pi)$ expected total cost-to-go when OB is taken
$CM_n(\pi)$ expected total cost-to-go when CM is taken
$g$ average cost per period
$b(\pi)$ bias under optimal policy
$b_{NA}(\pi)$ bias when NA is taken
$b_{PM}(\pi)$ bias when PM is taken
$b_{OB}(\pi)$ bias when OB is taken
I. INTRODUCTION

In many branches of industries, machines are operated under more or less stationary conditions. Most maintenance optimization models in the literature are based on the assumptions of these stationary conditions. However, wind turbines suffer from stochastic loadings. Wind speed varies season by season and day by day. Offshore turbines are also affected by stochastic wave climate. These stochastic loadings make the degradation process rather complex. In addition, the feasibility of conducting maintenance is also constrained by stochastic weather environments.

Current maintenance practice for wind farms mainly consists of scheduled maintenance and corrective (or breakdown) maintenance (CM). According to [1] and [2], scheduled maintenance is carried out usually twice a year for a single turbine, and there are on average 2.2 failures per turbine a year which require major repair. Considering today’s trend of large-scale wind farms with several tens or hundreds of turbines, and the long distance between the wind farm location and operation centers, the cost for these maintenance visits are substantial.

Thanks to the advancement of sensor technology, many turbine manufactures began to install condition-based monitoring (CBM) equipment with many sensors within turbines. It generates sensor signals to represent the turbine states. With the observed signals, one can presumably estimate the turbine’s physical condition, predict the failures, and finally make decisions about which maintenance actions should be taken. Consequently, wind farm operators can reduce the number of unnecessary visits and avoid unexpected, sometimes catastrophic, failures.

The CBM equipment provides abundant information, but it does not solve the uncertainty issue perfectly [3]. There are problems of measurement noises, and that a specific value of monitoring data could be coming from the different conditions of the target system. More importantly, that fault diagnosis based on sensor measurements is nontrivial because wind turbines operate under non-steady and irregular operating conditions. Oftentimes it is not possible to conclude the exact state of the turbine component. Instead, one will have to estimate the actual state in a probabilistic sense.

Three additional stochastic factors need to be considered for modeling wind turbine maintenance. The first factor is the weather conditions which may constrain the feasibility of maintenance actions. For example, under high wind speeds of more than 20 m/s, climbing up the turbines is not allowed. Under wind speeds higher than 30 m/s, the site becomes even inaccessible [4]. On the other hand, wind farms are inevitably located on windy sites to maximize electricity
generation. For this reason, repair actions cannot be carried out very often. In the preliminary study using a Monte-Carlo simulation performed by the Delft University of Technology, wind turbine availability remains only at 85%-94% in a 100 unit wind farm, situated about 35 kilometers off the Dutch coast [5]. The main reason of this relatively low availability is the farm’s poor accessibility, which is on average around 60%. In another study by Bussel [1], the availability of a wind farm was 76%.

The second factor is the repairing interruption and delay. Most wind farm-related repairs take several days to several weeks to complete. This relatively long repairing duration increases the likelihood that a repair is interrupted by adverse weather conditions. When the weather becomes adverse, the crew must stop working and wait until weather conditions become favorable. These delays cause revenue losses because wind turbines can no longer be operated until the repairs are completed.

The third factor to consider is the long lead time for assembling maintenance crews and attaining spare parts, which also significantly affects the downtime. For example, it can take several weeks for parts such as a gearbox to be delivered [6]. Due to the aforementioned uncertainty and stochastic issues, we believe that a properly timed and well-planned preventive maintenance strategy is pressingly needed in order to reduce the total maintenance costs and increase the availability of wind farms.

Taking all of the above issues into consideration, we derive the optimal repair strategies that minimize the average long-run cost under stochastic environments. We develop a multistate, dynamic optimization model, which considers several unique aspects of wind turbine operations, as discussed above. We emphasize the main contributions of this study with the following points.

1) We formulate the problem as a partially observed Markov decision process (POMDP), which considers the costs associated with different actions, the uncertainty of system states and weather effects. To the best of our knowledge, the proposed model is the first mathematical model for wind turbine maintenance.

2) We analytically derive the optimal control limits for each action as a set of closed-form expressions. We provide the necessary and sufficient conditions under which preventive maintenance will be optimal. The sufficient conditions for other actions to be optimal are also derived.

3) We establish several structural properties such as the monotonicity of the optimal policy.
We show that the structure of the optimal policy is similar to those studied in the previous POMDP literature, but our policy structure requires weaker assumptions. Optimality results for other policy structures, not previously proved in the literature, are also presented. We examine the practical implications of these properties in wind turbine maintenance.

The remainder of the paper is organized as follows. We start off with reviewing related work in Section II. Then, we present the POMDP model in Section III. In Section IV, several structural properties of the optimal policy are discussed. In Section V, we derive an algorithm for finding the optimal policy based on the structural properties established in the previous section. The computational results are reported in Section VI. Finally, we conclude the paper in Section VII with additional discussions and comments on future research.

II. LITERATURE REVIEW

Several studies have been conducted to find critical factors which affect the operations and maintenance (O&M) costs of wind turbine generators. Pacot et al. [6] discuss key performance indicators in wind farm management, and review the effects of several factors such as turbine age, turbine size, and location. Bussel [1] presents an expert system to determine the availability and O&M costs. The goal of Bussel’s study is to find the most economic solution by striking a balance between the front loading costs invested for reliability enhancement and the O&M costs. Rademakers et al. [5] further describe two simulation models for O&M and illustrate the features and benefits of their models through a case study of a 100 MW offshore wind farm.

Insightful review of the recent CBM for wind turbines are provided by Caselitz and Giebhardt [7]. The most widely used monitoring system is vibration monitoring. The other monitoring systems include, for example, measuring the temperature of bearings, lubrication oil particulate content analysis, and optical strain measurements [8]. Nilsson and Bertling [2] discuss the benefits of CBM with a case study of two wind farms. They present an asset life cycle cost analysis and break down the entire maintenance costs into several cost components. Mcmillan and Ault [4] also quantify the cost-effectiveness of CBM. They employ several probabilistic models to accommodate uncertainties, which are incorporated into a Monte Carlo simulation to capture the complex processes.

Several mathematical models have recently been introduced, which incorporate the information from CBM sensors. Although these models are not specifically developed for wind turbine
maintenance, we could gain insights regarding how CBM sensory information can be utilized. Maillart [9] uses POMDPs to adaptively schedule observations and to decide the appropriate maintenance actions based on the state information estimated from CBM sensory information. In this study, a system is assumed to undergo a multi-state Markovian deterioration process. Gebraeel [10] integrates the real-time sensory signals from CBM with a population-specific aging process in order to capture the degradation behavior of individual components. The author updates the remaining life distributions of individual components in a Bayesian manner. Similarly, Ghassemi et al. [11] represent system’s deterioration process with two sources. One is the average aging behavior that is usually provided by the manufacturer or estimated using survival data. The other source is the system utilization that can be diagnosed by using CBM data. They formulate the problem by a POMDP and derive optimal policies using dynamic programming.

Several studies examine the structural properties of POMDP maintenance models [9], [12], [13], [14], [15], [16]. Although these studies use different state definitions and cost structures, they establish a similar structural property of the optimal policy called the monotonic “At-Most-Four-Region” (AM4R) structure. The monotonic AM4R structure implies that along ordered subsets of deterioration state spaces, the optimal policy regions are divided into at most four regions with the following order: no action → observation (or, inspection) → no action → preventive maintenance. Ohnishi et al. [15] prove a similar result for the problem where a system is monitored incompletely in discrete decision epochs, but taking observation perfectly reveals the condition of a system at some cost. For detailed reviews of these AM4R studies, refer to [9].

Most maintenance studies in the literature consider static environmental conditions. Few quantitative studies have been done for systems operating under stochastic environments. Thomas et al. [17] investigate the repair strategies to maximize the expected survival time until a catastrophic event occurs in an uncertain environment. They consider the situation where a system should be stopped during inspection or maintenance action. If specific events, called “initiating events”, take place when a system is down or being replaced, it is denoted a catastrophic event. The examples can be found in military equipment or hospital systems. Consider a mission-critical military apparatus. If such an apparatus fails but a real dangerous situation happens, for instance, an adversary is attacking, the damage could be catastrophic. They show that similar AM4R structural results hold for a simple system where a system state takes only binary values,
i.e., operating or failed.

In this study we devise a multistate, POMDP model to represent the degradation process of wind turbines and to decide the optimal maintenance strategies. Our model extends the model introduced in [9] by incorporating several unique characteristics of wind turbine operations. In order to represent the stochastic weather conditions, our extended model adopts the “initiating events” idea, proposed in [17]. This is because the occurrence of harsh weather conditions delay repair processes and cause non-negligible revenue losses, making the circumstances analogous to those discussed in [17]. Other aspects of wind turbine operations such as the long lead time after an unplanned failure and the resulting production losses are also included in our model.

For our new wind turbine maintenance model, we show that the optimal decision rules are composed of the form of control limit polices. Previous POMDP studies in the literature show that there exists an optimal control limit for preventive maintenance action [15], [16]. However, they do not provide the exact value of the control limit. In this study, we derive closed-form conditions for each action to be optimal. These closed-form expressions help us determine the optimal strategies in an efficient way, leading to considerable computation reduction while handling high dimensional problems. We also show that our new model still holds the well-known monotonic AM4R policy structure under less strict assumptions than ones in [9]. Moreover, we demonstrate the conditions under which the structure becomes a more appealing monotonic “At-Most-Three-Region” (AM3R) structure, which has no second “no action” region.

III. MATHEMATICAL MODELS

In this section we formulate the wind turbine maintenance problem and introduce the existing algorithm to numerically solve it. In later sections, we will present a computationally improved algorithm after analyzing the structural properties of the presented model.

A. Model formulation

Let us consider a system whose deterioration levels are classified into a finite number of states $1, \cdots, m + 1$. State 1 denotes the best condition like “new”. State $m$ denotes the most deteriorated condition, and state $m + 1$ is the failed state. We define the state of a system as the following probability distribution

$$\pi = [\pi_1, \pi_2, \cdots, \pi_{m+1}],$$

(1)
where $\pi_i$, $i = 1, \ldots, m + 1$ is the probability that the system is in deterioration level $i$. $\pi$ is commonly known as an information state in the literature [16]. Since wind turbines do not operate upon failures, $\sum_{i=1}^{m} \pi_i = 1$ when the system is not failed. If it fails, $\pi_{m+1} = 1$. Since a state is defined as a probability distribution to represent the belief over the actual deteriorated condition, we formulate the problem using a POMDP model.

We assume that wind farm operators make decisions in discrete time. Conceptually, the duration of each decision period can be any time. But, we consider that decisions are made frequently because wind farm operators want to make a timely decision based on the stream of sensor signals from CBM equipment. For example, a weekly-based decision would be one of the practical choices. When decisions are made frequently, a discount rate is close to 1. Therefore, we formulate the problem as an average expected cost model as Puterman [18] suggests, and we examine the policies to minimize the expected cost per unit period.

Given the information state $\pi$, one of the following three actions is available for an operating system at the beginning of each period.

- **No Action (NA):** the action to continue the operation without any intervention. With this action, the system undergoes deterioration according to a known transition probability matrix, $P = [p_{ij}]_{(m+1) \times (m+1)}$. Suppose that the current information state is $\pi$ and no action is taken. The probability that the system will still operate until the next decision point is $R(\pi) = 1 - \sum_{i=1}^{m} \pi_i p_{i,m+1}$. People call this probability as the reliability of the system. Maillart [9] shows that the information state after the next transition, given the system is not failed, is

$$
\pi'_j(\pi) = \left\{ \begin{array}{ll}
\frac{\sum_{i=1}^{m} \pi_i p_{ij}}{R(\pi)}, & j = 1, 2, \ldots, m \\
0, & j = m + 1
\end{array} \right.
$$

As such, the system is transited to the next state $\pi'(\pi) = [\pi'_1(\pi), \ldots, \pi'_m(\pi), 0]$ with probability $R(\pi)$. If it fails with probability $1 - R(\pi)$, the state becomes $e_{m+1}$ in the next period.

- **Preventive Maintenance (PM):** the action to repair the system at cost $C_{PM} (\leq C_{CM})$, where $C_{CM}$ is the cost for $CM$. $PM$ takes one full period. We assume that in order to complete $PM$, weather conditions should be good during one full period. In the case that weather becomes harsh during $PM$, the crew hold the repair work until the weather returns to good conditions. This delay incurs $\tau$ revenue losses per period since wind turbines cannot be
operated until PM is completed. Let \( \bar{\pi} \) denote the state when the system is stopped for repairs. After PM, the state is returned to an as-good-as-new state.

- Observation (OB): the action to evaluate the exact deterioration level at cost \( C_{OB} \) (\( C_{OB} + C_{PM} \leq C_{CM} \)). OB instantaneously reveals the system state with certainty. So the information state reverts to state \( e_i \), where \( e_i = [0, \ldots, 1, \ldots, 0] \) is \( (m + 1) \times 1 \) dimensional row vector with a 1 in the \( i \)th position and 0 elsewhere. After observation, the decision maker will choose either NA or PM in that same decision period, based on the updated information state.

Upon a failure, parts are ordered and crews are arranged, which takes \( T \) lead time. When all of the parts and crew are available, we carry out CM for one full period at cost \( C_{CM} \) if weather conditions are good enough to do repair work. Otherwise, we have to wait until weather conditions are met. Unless CM is completed, wind turbines cannot be operated and it causes \( \tau \) revenue losses per period. After CM, the system is renewed to an as-good-as-new state, \( e_1 \). Figure 1 illustrates the process after a failure.

Let \( W_{PM} \) and \( W_{CM} \) represent the probabilities that adverse weather conditions, which prohibit PM or CM, occur in a period, respectively. When \( W_{PM} \) or \( W_{CM} > 0 \), it implies that the stochastic operating conditions affect maintenance feasibility. On the contrary, \( W_{PM} = W_{CM} = 0 \) represents a static environment in which repair actions can be taken any time. Since CM requires more complicated repair (or replace) jobs than PM, CM needs better weather conditions than PM. Therefore, we will have \( W_{CM} \geq W_{PM} \) in most practical cases. We assume that the events of these adverse weather conditions occur randomly.

Let \( V_n(\pi) \) denote the minimum expected total cost-to-go with \( n \) decision periods left when
the current state is $\pi$. We formulate the problem as the following average expected cost model:

$$V_n(\pi) = \min \left\{ NA_n(\pi) = (\tau T + CM_{n-T-1}(e_{m+1})) (1 - R(\pi)) + V_{n-1}(\pi') R(\pi), \\
PM_n(\pi) = (1 - W_{PM}) (\tau + C_{PM} + V_{n-1}(e_1)) + W_{PM} (\tau + PM_{n-1}(\tilde{\pi})) \\
OB_n(\pi) = C_{OB} + \sum_{i=1}^{n} Post_n(e_i) \pi_i \right\} \quad (3)$$

where

$$CM_{n-T-1}(e_{m+1}) = (1 - W_{CM}) (\tau + C_{CM} + V_{n-T-2}(e_1)) + W_{CM} (\tau + CM_{n-T-2}(e_{m+1})), \quad (4)$$

$$PM_{n-1}(\tilde{\pi}) = (1 - W_{PM}) (\tau + C_{PM} + V_{n-2}(e_1)) + W_{PM} (\tau + PM_{n-2}(\tilde{\pi})), \quad (5)$$

$$Post_n(e_i) = \min \{ NA_n(e_i), PM_n(e_i) \} \quad (6)$$

In (3), $NA_n(\pi)$, $PM_n(\pi)$ and $OB_n(\pi)$ are the total cost-to-go if we select $NA$, $PM$ and $OB$ at state $\pi$, respectively. In $NA_n(\pi)$, the first term $\tau T$ reflects revenue losses during lead time after a failure. $OB_n(\pi)$ and $Post_n(\pi)$ together represent that after each observation at cost $C_{OB}$, the state is updated to $e_i$ with probability $\pi_i$ and then we choose either $NA$ or $PM$ in the same decision period. $CM_{n-T-1}(e_{m+1})$ and $PM_{n-1}(\tilde{\pi})$ consider weather constraints when carrying out $CM$ or $PM$, respectively.

Since the system is renewed after $CM$ or $PM$ as long as weather conditions are good, the model is unichain for $0 \leq W_{CM} < 1$ and $0 \leq W_{PM} < 1$ [9]. That is, the transition matrix corresponding to each action consists of a single recurrent class. For these kinds of problems, Puterman [18] shows that $V_n(\pi)$ approaches a line with slope $g$ and intercept $b(\pi)$ as $n$ becomes large. That is,

$$\lim_{n \to \infty} V_n(\pi) = n \cdot g + b(\pi) \quad (7)$$

Here, $g$ denotes the average cost per unit time under the optimal policy, $b(\pi)$ is the bias, or the relative cost when the information state starts from $\pi$.

Taking the limits of both sides of (3), and then applying (7) in both sides yields

$$b(\pi) = \min \left\{ b_{NA}(\pi) = ((\tau - g)T + b(e_{m+1})) (1 - R(\pi)) + b(\pi') R(\pi) - g, \\
b_{PM}(\pi) = (1 - W_{PM}) (\tau + C_{PM}) + W_{PM} (b(\tilde{\pi}) + \tau) - g, \\
b_{OB}(\pi) = C_{OB} + \sum_{i=1}^{n} b(e_i) \pi_i \right\} \quad (8)$$
Applying the same technique to (4) and (5), respectively, yields,

\[ b(e_{m+1}) = C_{CM} + b(e_1) + \frac{\tau - g}{1 - W_{CM}} \]  \hspace{1cm} (9) \\
\[ b(\tilde{\pi}) = C_{PM} + b(e_1) + \frac{\tau - g}{1 - W_{PM}} \]  \hspace{1cm} (10)

Since \( b(\pi) \) is the relative difference in total cost that results from starting the process in state \( \pi \) instead of in any other state, Puterman [18] suggests to set \( b(\pi^0) = 0 \) for an arbitrary \( \pi^0 \).

Intuitively, we set \( b(e_1) = 0 \) in (9) and (10). Let us define the new maintenance costs which compound weather effects, lead time and production losses by \( C'_{CM} \) and \( C'_{PM} \), respectively, as follows:

\[ C'_{CM} = C_{CM} + \frac{\tau - g}{1 - W_{CM}} + (\tau - g)T \]  \hspace{1cm} (11) \\
\[ C'_{PM} = C_{PM} + \frac{\tau - g}{1 - W_{PM}} \]  \hspace{1cm} (12)

Note that both \( C'_{CM} \) and \( C'_{PM} \) are increasing in \( W_{CM} \) and \( W_{PM} \), respectively. This implies that higher frequency of harsh weather conditions incurs higher repair costs. Here, we assume that \( \tau \geq g \) because wind power system would not be profitable otherwise. Therefore, the added costs due to an unplanned failure (that is, \( C'_{CM} - C'_{PM} \)) arise from the following three factors: (1) increased repair costs (for doing \( CM \)) (2) increased possibility of repair delays due to more restricted weather requirements to carry out \( CM \) (3) production losses caused by the waiting time to prepare resources after a failure.

Plugging \( C'_{CM}, C'_{PM} \) simplifies (8)-(10) to

\[ b(\pi) = \min \left\{ \begin{array}{l}
    b_{NA}(\pi) = C'_{CM}(1 - R(\pi)) + b(\pi'(\pi))R(\pi) - g, \\
    b_{PM}(\pi) = C'_{PM}, \\
    b_{OB}(\pi) = C_{OB} + \sum_{i=1}^{m} b(e_i)\pi_i
\end{array} \right. \]  \hspace{1cm} (13)

**B. Solution method - pure recursive technique**

First, let us consider a sample path emanating from an information state \( \pi \). By a sample path, we mean the sequence of information states over time when no action is taken, which is denoted by \( \{\pi, \pi^2, \cdots, \Pi(\pi)\} \) where \( \pi^2 = \pi'(\pi), \pi^3 = \pi'(\pi^2) \) and so on. \( \Pi(\pi) \), defined by \( \Pi(\pi) \equiv \pi^{k*} \) where \( k* = \min\{k : ||\pi^{k+1} - \pi^k|| < \epsilon\} \) with small \( \epsilon \), is a stationary state or an absorbing state. Maillart [9] shows, by referring to [19], that when the Markov chain is acyclic, \( \Pi(\pi) \) exists for any \( \epsilon > 0 \).
Let us call the sequence of states emanating from one of the extreme points \( b(e_i), \forall i \), in (13) by an *extreme* sample path. Since all the biases at the states on the extreme sample paths are independent of the biases at the states on non-extreme sample paths, we can easily obtain \( b(e_i), \forall i \) and average cost \( g \) by applying policy iteration (or value iteration) methods to the states only on the extreme sample paths [18]. Then, \( b_{OB}(\pi) \) and \( b_{PM}(\pi) \) in (13) can be directly computed.

Now, we only need to compute \( b_{NA}(\pi) \) to get \( b(\pi) \). Maillart [20] introduces the following recursive technique. First, we solve (13) for \( \Pi(\pi) \) by

\[
\begin{align*}
    b(\Pi(\pi)) &= \min \left\{ b_{NA}(\Pi(\pi)) = C'_{CM} - \frac{g}{1 - R(\Pi(\pi))}, \\
    b_{PM}(\Pi(\pi)) = C'_{PM}, \\
    b_{OB}(\Pi(\pi)) = C_{OB} + \sum_{i=1}^{m} b(e_i) \Pi(\pi)_i \right\}
\end{align*}
\]

Then, we apply \( b(\Pi(\pi)) \) to (13) in order to find the optimal policy at the previous state. By solving the recursive set of equations backwards, we can get the optimal policy along the states on the sample path emanating from the original state \( \pi \).

However, this recursive technique might be computationally inefficient when we want to find the optimal policies at a large number of states in a high dimensional state space because we have to apply each recursive set of equations for each state. These computational difficulties motivate us to study further the structural properties of the model.

### IV. Structural properties

In this section we establish the several structural properties of the optimal policy. More specifically, we derive a set of closed expressions for the optimal policy including the exact control limits for PM. In later sections we will show how these results help attain optimal polices. We also present that the model exhibits the monotonous AM4R policy structure. This finding is an extension of a previous study in [9]. In [9] the AM4R results are shown, for a simper model than the presented model, under specific assumptions on the transition matrix and information states. We relax the assumptions to prove the results. We also establish the conditions where the optimal policy is simplified to the more intuitive AM3R structure.

#### A. Preliminary results

We first introduce several definitions which are often used in POMDP studies. Same definitions can be found in [12], [13], [15].
Definition 1: Information state $\pi$ is stochastically less (or smaller) than $\hat{\pi}$, denoted as $\pi \prec_{st} \hat{\pi}$ if and only if $\sum_{i=j}^{m+1} \pi_i \leq \sum_{i=j}^{m+1} \hat{\pi}_i$ for all $j = 1, \ldots, m + 1$.

Definition 2: Information state $\pi$ is less (or smaller) in likelihood ratio than $\hat{\pi}$, denoted as $\pi \prec_{lr} \hat{\pi}$ if and only if $\pi_i \hat{\pi}_j - \pi_j \hat{\pi}_i \geq 0$ for all $j \geq i$.

These two definitions present the binary relations of the two states in the sense of deterioration. Both definitions imply that when the system is less deteriorated, the state is stochastically (or in the likelihood ratio) less than another [16]. However, as Proposition 1(a) (see below) suggests, $\prec_{lr}$ relationship is stronger than $\prec_{st}$ relationship [13].

We also need additional definitions regarding the transition matrix $P$.

Definition 3: A transition matrix $P$ has an Increasing Failure Rate (IFR) if $\sum_{j \geq k} p_{ij} \leq \sum_{j \geq k} p_{i'j}$ for all $i' \geq i$ and $\forall k$.

Definition 4: A transition matrix $P$ is Totally Positive of order 2 (TP2) if $p_{ij}p_{i'j'} \geq p_{i'j}p_{ij'}$ for all $i' \geq i$ and $j' \geq j$.

These definitions imply that the more deteriorated system tends to more likely deteriorate further and/or fail [9]. Similar to the stochastic relations defined in Definition 1 and Definition 2, $TP2$ is more stringent assumption than $IFR$ due to the following Proposition 1(b) [13].

Proposition 1: (Rosenfield [13]) (a) If $\pi \prec_{lr} \hat{\pi}$, then $\pi \preceq_{st} \hat{\pi}$; (b) If $P$ is $TP2$, then $P$ is $IFR$.

Before presenting our results, we introduce several well-known results in the following two Propositions.

Proposition 2: (Derman [21]) For any column vector $v$ such that $v_i \leq v_{i+1}, \forall i$, if $\pi \prec_{st} \hat{\pi}$, then $\pi \cdot v \leq \hat{\pi}v$.

Proposition 3: (a) (Maillart [9]) Suppose that $P$ is IFR. If $\pi \prec_{st} \hat{\pi}$, then $R(\pi) \geq R(\hat{\pi})$. (b) (Maillart and Zheltova [16]) If $P$ is IFR and $\pi \prec_{st} \hat{\pi}$, then $\pi P \prec_{st} \hat{\pi}P$.
Now, Proposition 4 establishes that when $P$ is IFR, the stochastic ordering of two states are maintained after the transitions, and furthermore that when $P$ is upper-triangular, the next state is stochastically always larger than the current state. Upper-triangular transition matrix implies that a physical deteriorating condition cannot improve by itself, which represents natural deterioration behaviors of most practical systems. Proofs of all propositions, lemmas, and theorems are included in the appendix.

**Proposition 4:** (a) Suppose that $P$ is IFR. If $\pi \prec_{st} \hat{\pi}, \pi'(\pi) \prec_{st} \pi'(\hat{\pi})$. (b) Suppose that $P$ is IFR and that $P$ is upper-triangular. Then, $\pi \prec_{st} \pi'(\pi), \forall \pi$.

The following Proposition 5 demonstrates that the optimal cost-to-go for a failed system is always greater than, or equal to, the cost-to-go when it is stopped for $PM$.

**Proposition 5:** (a) $CM_n(e_{m+1}) - C_{CM} \geq PM_n(\pi) - C_{PM}$ for $\forall n$ where $CM_n(e_{m+1})$ and $PM_n(\pi)$ are defined in (4) and (3), respectively. (b) $CM_n(e_{m+1}) \geq PM_n(\pi)$ for all $n$.

The above Propositions allow us to derive the monotonicity of $V_n(\pi)$ in $\prec_{st}$-ordering, as shown in Lemma 1.

**Lemma 1:** If $P$ is IFR, $b(\pi)$ in (13) is non-decreasing in $\prec_{st}$ for all $n$.

The claim of Lemma 1 extends the result presented in [9] where she shows the monotonicity of the optimal cost function in $\prec_{lr}$-ordering on $TP2$ transition matrix. Also, the model in [9] is simpler than the presented model with the assumptions of $\tau = 0, T = 0$ and static environments (that is, $W_{CM} = W_{PM} = 0$). Therefore, the result of Lemma 1 is more general, and it can be applied to other general aging systems.

**B. Closed expressions for optimal policy regions**

In this section, we present the closed boundary expressions for the optimal policy. Let $\delta^*(\pi)$ denote the stationary optimal policy (or decision rule) at $\pi$. Also, let $\Omega_{NA}(\pi), \Omega_{OB}(\pi), \Omega_{PM}(\pi)$ be the set of information states with $\delta^*(\pi) = NA, \delta^*(\pi) = OB$ and $\delta^*(\pi) = PM$, respectively. In order to get the optimal policy to minimize the long-run average cost, we need to compare $b_{NA}(\pi), b_{PM}(\pi)$ and $b_{OB}(\pi)$. 
First, the following Lemma 2 explains when NA is preferred to PM and vice versa. To prove the claim, we apply the similar technique used in [11].

**Lemma 2:** Suppose that $P$ is IFR and upper-triangular. $\delta^*(\pi) \neq PM$ if $R(\pi) \geq 1 - \frac{g}{C_{CM} - C_{PM}}$. Also, $\delta^*(\pi) \neq NA$ if $R(\pi) < 1 - \frac{g}{C_{CM} - C_{PM}}$.

The claim of Lemma 2 is intuitive. As the system deteriorates, its reliability gets monotonically decreasing. When its reliability is lower than a threshold (here, it is $1 - \frac{g}{C_{CM} - C_{PM}}$), it is better to take some actions than doing nothing. On the contrary, we do not need to carry out costly maintenance action for a highly reliable system.

With the result of Lemma 2, $b_{OB}(\pi)$ in (13) can be reformulated as follows.

$$b_{OB}(\pi) = C_{OB} + \sum_{i=1}^{m} \{b_{NA}(e_i) \cdot I(R(e_i) \geq \alpha) + b_{PM}(e_i) \cdot I(R(e_i) < \alpha)\} \pi_i \quad (15)$$

Here, $I(\cdot)$ is the indicator function and $\alpha = 1 - \frac{g}{C_{CM} - C_{PM}}$. $OB_n(\pi)$ in (3) can be reformulated likewise.

Next, let us compare $b_{OB}(\pi)$ with $b_{PM}(\pi)$. If $C'_{PM} < C_{OB} + \sum b(e_i)\pi_i$, PM is preferred to OB. As a result, if $R(\pi) < 1 - \frac{g}{C_{CM} - C_{PM}}$ and $C'_{PM} < C_{OB} + \sum b(e_i)\pi_i$, the optimal policy is PM. Also, from the facts that $b_{OB}(\pi)$ is non-decreasing in $\prec_{st}$-ordering and that $b_{PM}(\pi)$ is constant, we can derive the control limit for PM in a closed form. Many previous maintenance studies based on a POMDP simply prove the “existence” of the control limit for PM. But, for this problem we analytically obtain the necessary and sufficient condition. Theorem 1 summarizes the results.

**Theorem 1:** Suppose that $P$ is IFR and upper-triangular. (a) The region where the optimal policy is PM is defined by $\Omega_{PM} = \{\pi; R(\pi) < 1 - \frac{g}{C_{CM} - C_{PM}}, C'_{PM} < C_{OB} + \sum b(e_i)\pi_i\}$, whereas PM cannot be optimal for $\pi \notin \Omega_{PM}$. (b) Furthermore, if $\delta^*(\pi) = PM$, $\delta^*(\hat{\pi}) = PM$ for $\pi \prec_{st} \hat{\pi}$.

This PM region in Theorem 1 defines the optimal PM region of the AM4R policy, as we will discuss in Section IV-D.

**Corollary 1:** Suppose that $P$ is IFR and upper-triangular. (a) If $R(\pi) < 1 - \frac{g}{C_{CM} - C_{PM}}$ and
\( C'_{PM} \geq C_{OB} + \sum b(e_i)\pi_i, \delta^*(\pi) = OB. \) (b) If \( R(\pi) \geq 1 - \frac{g}{C_{CM} - C_{PM}} \) and \( C'_{PM} < C_{OB} + \sum b(e_i)\pi_i, \delta^*(\pi) = NA. \)

Finally, let’s compare \( b_{NA}(\pi) \) with \( b_{OB}(\pi) \). We present the conditions under which \( NA \) is preferred to \( OB \) and vice versa, in Lemma 3 and Lemma 4.

**Lemma 3:** If \( R(\pi) \geq \frac{(C'_{CM} - C_{OB} - \sum b(e_i)\pi_i - g)}{C_{CM} - C_{OB} - \sum b(e_i)\pi_i^2}, \delta^*(\pi) \neq OB. \)

Similar to Lemma 2, Lemma 3 also explains that when the system is in a fairly good condition with a high reliability, we do not need to carry out costly inspection of the system. Along with Lemma 2, the following Corollary 2 specifies the sufficient condition for \( NA \) to be optimal; its proof follows directly from Lemma 2 and Lemma 3.

**Corollary 2:** If \( R(\pi) \geq \max\{1 - g - C_{CM} - C_{OB} - \sum b(e_i)\pi_i - g, C_{CM} - C_{OB} - \sum b(e_i)\pi_i - g\}, \delta^*(\pi) = NA. \)

Finally, Lemma 4 specifies the sufficient condition under which \( OB \) is optimal.

**Lemma 4:** Suppose that \( R(\pi) < \frac{C'_{CM} - C_{OB} - \sum b(e_i)\pi_i - g}{C_{CM} - C_{OB} - \sum b(e_i)\pi_i^2}. \) If \( \delta^*(\pi^2) = OB, \delta^*(\pi) = OB. \)

**C. Structural properties along sample path**

By extending the claim of Proposition 4(a), we can easily show that when \( P \) is IFR, all of the states in the sample path emanating from any \( \pi \) is in increasing stochastic order as long as \( \pi \prec \pi^2 \). Furthermore, if \( P \) is IFR and upper-triangular, \( \pi \prec_{st} \pi^2 \) always holds by Proposition 4(b) and thus, all states on any sample path is in \( \prec_{st} \)-increasing order. This allows us to apply all of the results developed in Section IV-B to the states along any sample path. The following Corollary 3 summarizes them.

**Corollary 3:** Suppose that \( P \) is IFR and upper-triangular. Then, the states along a sample path satisfy the following properties.

(a) Any sample path is in \( \prec_{st} \)-increasing order. That is, \( \pi \prec_{st} \pi^2 \prec_{st} \cdots \prec_{st} \Pi(\pi) \)
(b) \( V_n(\pi) \) and \( b(\pi) \) are non-decreasing along any sample path.
(c) Suppose that \( R(\pi^q) \geq 1 - \frac{g}{C_{CM} - C_{PM}} \). \( \delta^*(\pi^k) \neq PM \) for \( \forall k \leq q \). On the contrary, if \( R(\pi^q) < 1 - \frac{g}{C_{CM} - C_{PM}}, \delta^*(\pi^k) \neq NA \) for \( \forall k \geq q \).
(d) There exists a critical number \( k^* \) such that \( \delta^*(\pi^k) = PM, \forall k \geq k^* \), and \( \delta^*(\pi^k) \neq PM \) otherwise. And, such \( k^* \) is given by \( k^* = \max\{k1(\pi), k2(\pi)\} \) where

\[
\begin{align*}
  k1(\pi) &= \min\{k; R(\pi^k) < 1 - \frac{g}{C_{CM} - C'_{PM}}\} \quad (16) \\
  k2(\pi) &= \min\{k; C_{OB} + \sum b(e_i)\pi^k_i > C'_{PM}\} \quad (17)
\end{align*}
\]

### D. The monotonic At-Most-Four-Region policy

Several previous studies establish the AM4R policy structure along an ordered subset of state space for POMDP problems in different maintenance settings. For example, Maillart [9] presents the AM4R structure along any straight line of \( \prec_{it} \)-ordered information states when \( P \) is \( TP2 \) in her model.

In this section, we establish similar results for the presented problem under less stringent assumptions on the transition matrix and information states. More specifically, we show that the optimal policy has the AM4R structure along a straight line of \( \prec_{st} \)-ordered states on \( IFR \) transition matrix. Consider two states \( \pi \) and \( \hat{\pi} \) for \( \pi \prec_{st} \hat{\pi} \). Let’s denote a state between \( \pi \) and \( \hat{\pi} \) by \( \pi(\lambda) = \lambda\pi + (1-\lambda)\hat{\pi}, 0 \leq \lambda \leq 1 \). Here, higher \( \lambda \) implies a more deteriorated condition (we will show the reason in Theorem 2). Then, there exist at most three numbers \( \lambda_1, \lambda_2, \lambda_3 \) to divide the optimal policy regions as follows.

\[
\delta^*(\pi(\lambda)) = \min\begin{cases} 
  NA, & \text{if } \lambda < \lambda_1 \text{ or } \lambda_2 < \lambda \leq \lambda_3 \\
  OB, & \text{if } \lambda_1 \leq \lambda \leq \lambda_2 \\
  PM, & \text{if } \lambda > \lambda_3 
\end{cases} \quad (18)
\]

That is, as \( \lambda \) increases, the optimal policy regions are divided into at most four regions with the order \( NA \rightarrow OB \rightarrow NA \rightarrow PM \). To establish this AM4R structure, we first show the concavity of \( V_n(\pi) \).

**Lemma 5:** \( V_n(\pi) \) is piecewise linear concave for all \( n \).

Now, we are ready to prove the monotonic AM4R structure along a \( \prec_{st} \)-increasing line.

**Theorem 2:** IF \( P \) is IFR, the optimal policy has the monotonic AM4R structure along any straight line of information states in \( \prec_{st} \)-increasing order. Furthermore, the control limit to
define optimal PM policy is defined by
\[
\lambda^* = \inf\{\lambda; R(\pi(\lambda)) < 1 - \frac{g}{C_{CM} - C_{PM}'}, C_{PM}' < C_{OB} + \sum b(e_i)\pi(\lambda)_i\}.\]

As Rosenfield [13] points out, the second NA region in the AM4R structure may seem counter-intuitive. In the following discussions, we establish the conditions under which we have the more intuitive AM3R policy structure. Let’s define the critical numbers to divide the optimal policy regions as follows.

\[
\lambda_{NA \leq PM} = \max\{\lambda; R(\pi(\lambda)) \geq 1 - \frac{g}{C_{CM} - C_{PM}'}\}
\]

(19)

\[
\lambda_{OB \leq PM} = \max\{\lambda; C_{OB} + \sum b(e_i)\pi(\lambda)_i \geq C_{PM}'\}
\]

(20)

Note that for \(\lambda \leq \lambda_{NA \leq PM}\), NA is preferred to PM and vise versa. Similarly, For \(\lambda \leq \lambda_{OB \leq PM}\), OB is preferred to PM and vise versa.

Corollary 4: If \(\lambda_{NA \leq PM} < \lambda_{OB \leq PM}\), the optimal policy has the monotonic AM3R structure along any \(\prec_{st}\)-increasing straight line of information states with the order of \(NA \rightarrow OB \rightarrow PM\). The optimal policy regions for PM is given by \(\{\pi(\lambda); C_{PM}' < C_{OB} + \sum b(e_i)\pi(\lambda)_i\}\).

Figure 2 compares the two policy structures. Whether the optimal policy structure exhibits the AM4R or AM3R is highly dependent on the costs of PM. When the preventive repairing costs are relatively high compared to the costs related to other actions, the structure is more likely to result in the AM4R structure, as shown in Figure 2(a). Otherwise, when PM costs are comparable to other costs, it is more likely ended with the AM3R structure, as shown in Figure 2(b).

In wind turbine operations, the repair costs \((C_{CM}')\) after an unplanned failure are considerably expensive compared to the PM cost \((C_{PM}')\), as explained in Section I and Section III. Also, in most cases OB costs are not negligible because inspecting the physical condition by dispatching crew is costly due to the high labor costs and the long distance of wind farms from the operation centers [22]. This implies that the presented optimal policy would more likely lead to the AM3R structure in most of wind turbine maintenance problems.
V. ALGORITHM

In Section III-B, we introduced the pure recursive technique to get optimal policy. Now, using the structural policies developed so far, we can reduce the computational efforts substantially.

First, given the parameter values \((C_{CM}, C_{PM}, C_{OB}, w, P, \tau \text{ and } T)\), we obtain \(b(e_i), i = 1, \cdots, m + 1\) and average cost \(g\) by applying policy (or value) iteration to the states on the extreme sample paths. Then, \(C'_{CM}, C'_{PM}\) in (11) and (12) can be computed, respectively, as discussed in Section III-A. Then, we apply the following decision rules to attain the optimal policy for a given \(\pi\).

- Suppose that there exists a \(\pi\) at which the optimal policy is \(PM\). \(\delta^*(\tilde{\pi}) = PM\) for \(\pi \prec_{st} \tilde{\pi}, \forall \tilde{\pi}\).
Suppose that $R(\pi) < 1 - \frac{g}{C_{CM} - C_{PM}}$. If $b_{OB}(\pi) > b_{PM}(\pi)$, $\delta^*(\pi) = PM$. Otherwise, $\delta^*(\pi) = OB$.

Suppose that $R(\pi) \geq 1 - \frac{g}{C_{CM} - C_{PM}}$. $\delta^*(\pi) = NA$ if $b_{OB}(\pi) > b_{PM}(\pi)$ or if $R(\pi) \geq \frac{C'_{CM} - C_{OB} - \sum b(e_i)\pi_i - g}{C_{CM} - C_{OB} - \sum b(e_i)\pi_i}$.

Suppose that $1 - \frac{g}{C_{CM} - C_{PM}} \leq R(\pi) < \frac{C'_{CM} - C_{OB} - \sum b(e_i)\pi_i - g}{C_{CM} - C_{OB} - \sum b(e_i)\pi_i}$ and $b_{OB}(\pi) \leq b_{PM}$. We apply the following recursive method, which improves the pure recursive technique.

- Step 1. Set $k = 1$;
- Step 2. If $R(\pi^k) < 1 - \frac{g}{C_{CM} - C_{PM}}$, $b(\pi^k) = \min\{b_{OB}(\pi^k), b_{PM}(\pi^k)\}$. Then apply the recursive set of equations (13) backward to get $b(\pi^{k-1}), \ldots, b(\pi)$. Otherwise, $k = k + 1$ and go to Step 3.
- Step 3. If $||\pi^{k+1} - \pi^k|| < \epsilon$, we apply (14) to get $b(\pi^k)$ and then step backwards along the path by comparing $b_{NA}$ and $b_{OB}$ to get $b(\pi^{k-1}), \ldots, b(\pi)$. Otherwise, $k = k + 1$ and go back to Step 2.

The above method means that in many cases, optimal policy can be analytically obtained from the closed expressions. We need to apply the recursive method only for the states whose reliabilities are between $1 - \frac{g}{C_{CM} - C_{PM}}$ and $\frac{C'_{CM} - C_{OB} - \sum b(e_i)\pi_i - g}{C_{CM} - C_{OB} - \sum b(e_i)\pi_i}$. Even for the recursive method itself, as Step 2 shows, we do not have to proceed until we meet the stationary state $\Pi(\pi)$. Along the sample path, once we find the state whose optimal policy is not $NA$ (That is, $R(\pi^k) \leq 1 - \frac{g}{C_{CM} - C_{PM}}$ for some $\pi^k$), we can compute $b(\pi^k)$ by comparing $b_{PM}$ with $b_{OB}(\pi^k)$. Then we can step backwards by applying (13) until we get $\pi$. On the contrary, Step 3 occurs when the reliability at the stationary state is greater than $1 - \frac{g}{C_{CM} - C_{PM}}$. In this case, $PM$ cannot be optimal at all of the states along the sample path originating from $\pi$. Therefore, we only need to compare $b_{NA}(\pi^k)$ with $b_{OB}(\pi^k)$ when we step backwards to $\pi$.

VI. NUMERICAL EXAMPLES

McMillan and Ault [4] show that the most critical failures are associated with a gearbox because of high capital cost, long lead time for repairs, difficulty in replacing a gearbox, and lengthy downtime compounded by adverse weather conditions. Therefore, we choose a gearbox among several components of a wind turbine to illustrate the presented model.
A. Example for gearbox maintenance

We choose appropriate parameter values based on the published data and discussions with our industry partners. For the costs for repairing a gearbox, we refer to [22]. The total direct costs for $CM$, which include labor costs, crane rent, materials and consumables, are $C_{CM} = 12,720$. The $PM$ costs are about a half of $CM$ costs, that is, $C_{PM} = 6,360$. For a 2.5 MW turbine, revenue loss during one week is $\tau = 8,820$. We set $C_{OB} = 100$, which is about 10% of $C_{CM}$, according to the suggestions of our industry partners (The monetary unit of each cost factor in this example is euro).

Typical downtime after failures may take 600 hours (25 days) up to 60 days [4], [8]. The major contribution of this lengthy down time is the long lead time when the spare parts and/or crew are not available. In this study, we assume that upon failure, the lead time ($T$) for assembling repair crew and spare parts and travel time takes six weeks. Repair can be carried out in about one week [4].

Generally, a transition matrix $P$ can be generated from historical data by taking a long-run history about the deterioration states and counting transitions. Due to the relatively short history of preventive maintenance practices in wind turbine industries, we do not yet have a transition matrix generated from an actual aging gearbox. So we apply similar matrix used in the example in [9] with slight modifications (in the next section, we will examine the sensitivity of $P$). We assume that the weekly-based deterioration process follows Markovian behaviors with the following $IFR$, uppertriangular $P$ matrix:

$$ P = \begin{bmatrix}
0.90 & 0.05 & 0.03 & 0.02 \\
0.00 & 0.85 & 0.10 & 0.05 \\
0.00 & 0.00 & 0.92 & 0.08 \\
0.00 & 0.00 & 0.00 & 1.00
\end{bmatrix} \quad (21) $$

The state can be represented as a four dimensional row vector, $\pi = \{\pi_1, \pi_2, \pi_3, \pi_4\}$. $\pi_1$, $\pi_2$ and $\pi_3$ represent the probabilities of being in a normal, alert and alarm condition, respectively. Note that $\pi_1 + \pi_2 + \pi_3 = 1$ for operating turbines.

Figure 3 illustrates the optimal policies with two different stochastic weather environments. We can see that $\Omega_{OB}$ and $\Omega_{PM}$ are convex sets. Also, if we draw a line between any two points, the policy regions are divided into at most three regions, which is consistent with the previous discussions that the AM3R structure might dominate over the AM4R structure in most
real applications. It is also noticeable that $\Omega_{PM}$ gets smaller as the chance of adverse weather conditions to prohibit $PM$ increases. That is, with higher frequency of adverse weather conditions (that is, with higher $W_{PM}$), wind farm operators should be more conservative in carrying out $PM$ because of possible production losses caused by interrupted and delayed jobs during harsh weather.

![Diagram](image)

Fig. 3. Optimal policies

Figure 4 superimposes the control limits developed in Section IV-B on the optimal policy for the same example in Figure 3.(a). Line 1 depicts the preference of $NA$ to $PM$ or vise versa with the equation $R(\pi) = 1 - \frac{g}{c'_{CM}-c'_{PM}}$. Line 2 is obtained from the comparison of $b_{OB}$ and $b_{PM}$ with the equation $C'_{PM} = C_{OB} + \sum b(e_i)\pi_i$. Finally, Line 3 defines the area where $NA$ is
preferred to \( OB \). The optimal policy of each area is as follows.

- \( PM \) in states above Line 1 and Line 2 (by Theorem 1)
- \( OB \) in states above Line 1 and below Line 2 (by Corollary 1(a))
- \( NA \) in states below Line 1 and Line 3 by (Lemma 2 and Lemma 3)
- \( NA \) in states in the triangular area surrounded by Line 1, Line 2 and Line 3 (by Corollary 1(b))

The only states whose optimal policy are not straightforward from these control limits are shown in the shaded region surrounded by the dashed lines in Figure 4. The optimal policy in this region is obtained by applying the improved recursive technique discussed in Section V.

B. Performance comparison

Suppose that we want to find the optimal policy at every grid point as shown in the previous figures in Section VI-A. As a dimension of states increases, computation time significantly increases when we use the pure recursive algorithm. Figure 5 compares the performance of the suggested algorithm with the pure recursive technique. We use the same parameter values in the previous examples and \( W_{PM} = 0.1 \) and \( W_{CM} = 0.4 \), but vary the transition matrix along a state size. The results indicate that the closed forms of decision boundaries compounded by the improved recursive technique reduce the computation time by more than 70% over various sizes of the problems.
C. Sensitivity analysis of transition matrix \( P \)

Considering difficulties to get a transition matrix \( P \), we analyze the sensitivity of a transition matrix by applying additional four different matrices, \( P_i, i = 1, \cdots, 4 \). \( P_1 \) represents a more slowly deteriorating system in a stochastical sense than \( P \) in (21). That is, each row vector of \( P_1 \) is stochastically less than the corresponding row vector of \( P \). Let’s denote this relationship by \( P_1 \prec_{st} P \). Similarly, \( P_2 \prec_{st} P \) and also \( P_1 \prec_{st} P_2 \). On the other hand, \( P_3 \) and \( P_4 \) represent more rapidly deteriorating systems than \( P \), such that \( P \prec_{st} P_3 \prec_{st} P_4 \).

We quantify the speeds of deterioration of a system with \( P_k, k = 1, \cdots, 4 \), compared to \( P \), with the following measure.

\[
\Delta P_k(\%) = \sum_{i=1}^{m} \sum_{j \geq i} \frac{|P(i, j) - P_k(i, j)|}{P(i, j)} \times 100, \tag{22}
\]

where \( P(i, j) \) is the element in \( i^{th} \) row and \( j^{th} \) column of \( P \) matrix and \( P_k(i, j) \) is similarly defined. Note that the lower off-diagonal elements are not involved in (22) because we consider upper-triangular matrices. \( \Delta P_k \) implies the relative difference of \( P_k \), compared to \( P \).

To measure the sensitivity of a transition matrix, we use simulations. Suppose that the actual system undergoes deterioration process following a transition matrix \( P \). We simulate the trajectories of system states following \( P \) from 136 different starting points. Here, 136 starting points are the points in the grid, similar to the grid points shown in Figure 4. But, to speed up the simulations, we use a coarser grid such that the distance between adjacent grid points is \( \frac{2}{3} \). From each starting point, the simulation is performed over 1,000 periods. At each period, we
take actions as the optimal policy suggests. Then the costs are averaged. This process is repeated 30 times. That is, we gain the average cost $g$ by the simulations on 136 different starting points $\times$ 1,000 periods $\times$ 30 trajectories (runs).

However, suppose that we do not know the transition matrix exactly, so we incorrectly use the transition matrix $P_k$ to attain optimal policies, while the actual deterioration process follows $P$. We apply the similar simulation process, but we use $P_k$ to decide the optimal policy. Then, we compute the average cost $g_k$.

From the results of the simulations, we quantitatively measure the sensitivity of each transition matrix by

$$\Delta G_k = \frac{g_k - g}{g} \times 100.$$  \hspace{1cm} (23)

Table I summarizes the sensitivity results. The fourth column (that is, $g_k$) shows that the average costs increase as the assumed transition matrix $P_k$ deviates from the actual transition matrix $P$. However, the difference is not significant as the fifth column (that is, $\Delta G_k$) indicates. Even when the values of actual transition matrix deviates from the assumed transition matrix values by about 10% such as $P_1$ and $P_4$, the increased cost is about 2.0% on average. When the element values are different by 5-6% such as $P_2$ and $P_3$, average costs are increased by around 1%.

Although the results show that the average costs are not seriously affected by the deviation of the assumed transition matrix from the actual one, we recommend that considerable efforts to accumulate data regarding system deterioration are necessary. Rademakers et al. [22] also suggest that industry parties should collaborate with each other to collect and share data for the improvement of O&M of wind turbines. For conventional power systems, these data for critical equipment such as circuit breakers and transformers have been accumulated and several preventive maintenance strategies have been introduced based on historical data [23], [24], [25]. Similar efforts are necessary in wind power industries.

VII. Summary

Despite the vastly potential capacity of wind power reserve, the share of wind energy still remains a small portion of the entire energy market. One of the critical factors for enhancing the
marketability of wind energy is to cut the O&M costs because dispatching maintenance crew with heavy-duty equipment to a remote wind farm site is very costly.

In this study we use probabilistic cost modeling to quantify risks and uncertainties to develop optimal maintenance decision models for the O&M of wind turbines. The O&M of wind turbines have unique aspects to consider such as stochastic weather conditions that affect repair work, long lead time to prepare maintenance crew and materials after failures, and the significant revenue losses. These unique aspects are incorporated into the model.

We analytically derive a set of closed-form expressions for the optimal policy and show how these results can be utilized to solve large size problems. We also extend the AM4R structure under weaker assumptions than the previous study in the literature, and also demonstrate the conditions under which this AM4R structure becomes the AM3R structure. These results should be applicable to other general aging systems.

As future work, we could extend the model to incorporate multiple wind turbines. In this study we assume independence of each turbine operation. However, when turbines are operating, the rotating blades change wind speeds, which affect other turbine’s operation; this is known as “wake effects”. It would be interesting to see how robust the recommended maintenance policy can perform in a wind farm that houses many wind turbines. As a part of our ongoing research, we are developing a large-scale simulation model for wind farm operations using the discrete event specification (DEVS) formalism [26] with hundreds of wind turbines [27]. We will integrate the DEVS simulation model with the presented model to validate the optimal policy, and to see if further modifications are necessary when we consider multiple turbines.

Another direction for future work is to extend the model such that the parameters represent non-homogeneous, time-varying behaviors of wind turbine operations. For example, weather conditions can exhibit more complicated situations. The presented model assumes that harsh
weather conditions take place randomly. In the future, we might consider weather environments with different forecasting capabilities. The structural properties developed in this study will be exploited to solve these problems with more complex settings.

APPENDIX

Proof of Proposition 4:
(a) Let \( \pi_i \) and \( (\pi P)_i \) denote the \( i \)th position of the row vector \( \pi \)\( = \pi'(\pi) \) and \( \pi P \), respectively. Then, we have

\[
\sum_{i \geq j} \pi_i^2 = \sum_{i \geq j} \frac{(\pi P)_i}{R(\pi)} \leq \sum_{i \geq j} \frac{\pi_k}{R(\pi)} = \sum_{i \geq j} \pi_i^2
\]

(24)

The two inequalities in (24) hold due to Proposition 3(a) and Proposition 3(b), respectively.

(b)

\[
\sum_{j \geq k} \pi_j^2 = \sum_{j \geq k} \frac{\pi_i P_{ij}}{R(\pi)} \geq \sum_{i \geq j} \sum_{j \geq k} \pi_i P_{ij}
\]

\[
= \sum_{i \geq k} \pi_i \sum_{j \geq k} P_{ij} \geq \sum_{i \geq k} \pi_i \sum_{j \geq k} P_{ij}
\]

\[
= \sum_{i \geq k} \pi_i
\]

(25)

The last equality in (25) is from the assumption of having an upper-triangular \( P \).

Proof of Proposition 5:
(a) We prove the claim by induction method. Without loss of generality, suppose that \( CM_0(e_{m+1}) = C_{CM} \) and \( PM_0(\pi) = C_{PM} \). Then, \( CM_1(e_{m+1}) - C_{CM} = PM_n(\pi) - C_{PM} = \tau \). Suppose that \( CM_n(e_{m+1}) - C_{CM} \geq PM_n(\pi) - C_{PM} \). Then,

\[
CM_{n+1}(e_{m+1}) - C_{CM}
\]

\[
= (1 - W_{CM})(\tau + C_{CM} + V_n(e_1)) + W_{CM}(\tau + CM_n(e_{m+1}) - C_{CM})
\]

(26)

\[
= \tau + V_n(e_1) + W_{CM}(CM_n(e_{m+1}) - C_{CM} - V_n(e_1))
\]

(27)

\[
\geq \tau + V_n(e_1) + W_{PM}(PM_n(e_{m+1}) - C_{PM} - V_n(e_1))
\]

(28)

\[
= PM_{n+1}(\pi) - C_{PM},
\]

(29)

where (28) is from induction hypothesis. Therefore, \( CM_n(e_{m+1}) - C_{CM} \geq PM_n(\pi) - C_{PM} \) holds for \( \forall n \).
Consider the following POMDP model which is similar to the presented model in (3) to (6), but without lead time

**Proof of Lemma 1:**

Consider the following POMDP model which is similar to the presented model in (3) to (6), but without lead time to prepare resources (that is, \( T = 0 \)).

\[
J_n(\pi) = \min \left\{ \begin{array}{ll}
NA^J_n(\pi) = CM^J_n(e_m+1)(1 - R(\pi)) + J_{n-1}(\pi^2)R(\pi) \\
PM^J_n(\pi) = (1 - W_{CM})(\tau + C_{PM} + J_{n-1}(e_1)) + W_{PM}(\tau + PM_{n-1}(\pi)) \\
OB^J_n(\pi) = C_{OB} + \sum_{i=1}^{\infty} Post^J_n(e_i)\pi_i
\end{array} \right.
\]

where \( CM^J = CM + \tau T, C^J_{CM} = C_{PM}, \) and \( C^J_{OB} = C_{OB} \). Also, \( CM^J_{n-1}(\cdot), PM^J_{n-1}(\cdot) \) and \( Post^J_n(\cdot) \) are defined likewise as in (4) to (6).

By induction, we can show that \( J_n(\pi) \) is non-decreasing in \( \prec_{st} \) for all \( n \) when \( P \) is IFR. Suppose that \( CM_0(e_m+1) = CM^J, PM^J_0(\pi) = C^J_{PM} \) where \( CM^J \) and \( C^J_{PM} \) are corrective and preventive maintenance costs, respectively. Also, suppose that \( J_0(\pi) = 0, \forall \pi \) for an operating system. Then, \( NA^J_1(\pi) = C^J_{CM}(1 - R(\pi)) \) is non-decreasing in \( \prec_{st} \) from Proposition 3(a), and \( PM_1(\pi) = \tau + C^J_{PM} \) is constant. \( OB_1(\pi) = C_{OB} + \sum_{i=1}^{\infty} min\{NA^J_i(e_i), PM^J_i(e_i)\}\pi_i \). Since \( e_i \prec_{st} e_j \) for \( i \leq j \), \( NA^J_i(e_i) \) is non-decreasing in \( i \). Thus, \( OB_1(\pi) \) is also non-decreasing in \( \prec_{st} \) due to Proposition 2. Therefore, \( J_1(\pi) \) is non-decreasing in \( \prec_{st} \). Suppose that \( J_n(\pi) \) is non-decreasing in \( \prec_{st} \). Then, for \( \pi \prec \tilde{\pi} \),

\[
NA^J_{n+1}(\pi) = CM^J_n(e_m+1)(1 - R(\pi)) + J_n(\tilde{\pi}^2)R(\pi)
\]

(36) follows from the induction assumption and Proposition 4(a). (38) follows from Proposition 3(a), Proposition 5 and the fact that \( J_n(\pi) \leq PM^J_n(\pi) \leq CM^J_n(e_m+1) \) for \( \forall \pi \). It is obvious that \( OB^J_n(\pi) = C^J_{OB} + \sum_{i=1}^{\infty} min\{NA^J_{n+1}(e_i), PM^J_{n+1}(e_i)\}\pi_i \) is also non-decreasing in \( \prec_{st} \) with the similar reason explained above. Consequently, \( J_{n+1}(\pi) \) is nondecreasing in \( \prec_{st}, \forall n \).
$J_n(\pi)$ approaches a line with slope $g^J$ and intercept $b^J(\pi)$ as $n$ becomes large. That is, $\lim_{n \to \infty} J_n(\pi) = n \cdot g^J + b^J(\pi)$. By taking limits in both sides of (34) and applying the similar technique introduced in Section III, we have the following $b^J(\pi)$, which is similar to $b(\pi)$ in (13).

$$b^J(\pi) = \min \left\{ \begin{array}{l}
  b^J_{NA}(\pi) = C^J_{CM}(1 - R(\pi)) + b^J(\pi' \pi)R(\pi) - g^J, \\
  b^J_{PM}(\pi) = C^J_{PM}, \\
  b^J_{OB}(\pi) = C^J_{OB} + \sum_{i=1}^{m} b^J(e_i)\pi_i
\end{array} \right. \tag{40}$$

where $C^J_{CM}$ and $C^J_{PM}$ can be obtained likewise as in (11), (12), respectively after setting $T = 0$ in (11). If we set $g = g^J \left( \frac{2W_{CM} - R(\pi)}{1-W_{CM}+(1-R(\pi))(1+C_{CM})T} \right)$, $b^J(\pi)$ and $b(\pi)$ are equivalent. Since $J_n(\pi)$ is non-decreasing in $\prec_{st}$ for all $n$, so are $b^J(\pi)$ and $b(\pi)$, which concludes the claim. $lacksquare$

**Proof of Lemma 2:**

$$b_{NA}(\pi) - b_{PM}(\pi) = C^J_{CM}(1 - R(\pi)) + b(\pi^2)R(\pi) - g - C^J_{PM} \tag{41}$$

$$= (C^J_{CM} - C^J_{PM})(1 - R(\pi)) - g + (b(\pi^2) - C^J_{PM})R(\pi) \tag{42}$$

Note that $b(\pi^2) \leq C^J_{PM}$. Consequently, if $(C^J_{CM} - C^J_{PM})(1 - R(\pi)) - g \leq 0$ (or equivalently, $R(\pi) \geq 1 - \frac{g}{C^J_{CM} - C^J_{PM}}$), $NA$ is preferred to $PM$.

Next, consider the case that $(C^J_{CM} - C^J_{PM})(1 - R(\pi)) - g > 0$. Let’s assume that $\delta^*(\pi) = NA$. Then,

$$b(\pi^2) - b(\pi) = b(\pi^2) - (C^J_{CM}(1 - R(\pi)) + b(\pi^2)R(\pi) - g) \tag{43}$$

$$= (b(\pi^2) - C^J_{PM})(1 - R(\pi)) - (C^J_{CM} - C^J_{PM})(1 - R(\pi)) + g \tag{44}$$

(43) holds from the assumption $\delta^*(\pi) = NA$ and thus, $b(\pi) = C^J_{CM}(1 - R(\pi)) + b(\pi^2)R(\pi) - g$. Note that in (44), $b(\pi^2) \leq C^J_{PM}$. Therefore, when $(C^J_{CM} - C^J_{PM})(1 - R(\pi)) - g > 0$, $b(\pi^2) \leq b(\pi)$ with the assumption of $\delta^*(\pi) = NA$. But, this result contradicts that $b(\pi^2) \geq b(\pi)$ for all $\pi$ from Proposition 4(b) and Lemma 1. Therefore, when $(C^J_{CM} - C^J_{PM})(1 - R(\pi)) - g < 0$, or equivalently, $R(\pi) < 1 - \frac{g}{C^J_{CM} - C^J_{PM}}$, $NA$ cannot be the optimal action. $lacksquare$

**Proof of Theorem 1:**

The first part is straightforward from Lemma 2 and the above discussions. Regarding the second part, $NA$ cannot be optimal at $\hat{\pi}$ from the fact that $R(\hat{\pi}) \leq R(\pi)$ for $\pi \prec_{st} \hat{\pi}$. Also, since $b(e_i)$ is non-decreasing in $i$, $\sum_i b(e_i)\pi_i$ is also non-decreasing in $\prec_{st}$-ordering from Proposition 2, and so is $b_{OB}(\pi)$. This leads to $b_{OB}(\hat{\pi}) \geq b_{OB}(\pi)$. But, $b_{PM}(\pi)$ is constant. Thus, when $\delta^*(\pi) = PM$, $OB$ cannot be optimal at $\hat{\pi}$ as well, which concludes the second part of the Theorem. $lacksquare$

**Proof of Corollary 1:**

It follows directly from Lemma 2 and the fact that $OB$ is preferred to $PM$ when $C^J_{PM} \geq C_{OB} + \sum b(e_i)\pi_i$. $lacksquare$

**Proof of Lemma 3:**
We use similar technique used in Lemma 2.

\[ b_{NA}(\pi) - b_{OB}(\pi) \]
\[ = C'_{CM}(1 - R(\pi)) + b(\pi^2)R(\pi) - g - C_{OB} - \sum b(e_i)\pi_i \]  \hspace{1cm} (45)
\[ = (C'_{CM} - C_{OB} - \sum b(e_i)\pi_i)(1 - R(\pi)) - g + R(\pi)(b(\pi^2) - C_{OB} - \sum b(e_i)\pi_i) \]  \hspace{1cm} (46)
\[ = (C'_{CM} - C_{OB} - \sum b(e_i)\pi_i)(1 - R(\pi)) - g + R(\pi)\sum b(e_i)(\pi_i^2 - \pi_i) \]
\[ + R(\pi)(b(\pi^2) - C_{OB} - \sum b(e_i)\pi_i^2), \]  \hspace{1cm} (47)

Note that \( b(\pi^2) \leq C_{OB} + \sum b(e_i)\pi_i^2 \). Therefore, if \( (C'_{CM} - C_{OB} - \sum b(e_i)\pi_i)(1 - R(\pi)) - g + R(\pi)\sum b(e_i)(\pi_i^2 - \pi_i) \leq 0 \), \( b_{NA}(\pi) \leq b_{OB}(\pi) \). Re-arranging the condition yields

\[ (C'_{CM} - C_{OB} - \sum b(e_i)\pi_i)(1 - R(\pi)) - g + R(\pi)\sum b(e_i)(\pi_i^2 - \pi_i) < 0 \]  \hspace{1cm} (48)
\[ \Leftrightarrow R(\pi) \geq \frac{C'_{CM} - C_{OB} - \sum b(e_i)\pi_i - g}{C'_{CM} - C_{OB} - \sum b(e_i)\pi_i^2} \]  \hspace{1cm} (49)

The last inequality (50) comes from \( b(e_i) \leq C'_{PM} \) for all \( i = 1, \cdots, m \) and from \( C_{OB} + C'_{PM} \leq C_{CM} \) (Note that \( C_{OB} + C_{PM} \leq C_{CM} \) by assumption).

**Proof of Corollary 2:**

It follows directly from Lemma 2 and Lemma 3.

**Proof of Lemma 4:**

We will use contradiction. Assume that \( \delta^*(\pi) = NA \). Then,

\[ b(\pi^2) - b(\pi) - \sum b(e_i)(\pi_i^2 - \pi_i) \]  \hspace{1cm} (51)
\[ = b(\pi^2) - C'_{CM}(1 - R(\pi)) - b(\pi^2)R(\pi) + g - \sum b(e_i)(\pi_i^2 - \pi_i) \]  \hspace{1cm} (52)
\[ = (b(\pi^2) - C'_{CM})(1 - R(\pi)) + g - \sum b(e_i)(\pi_i^2 - \pi_i) \]  \hspace{1cm} (53)
\[ = (b(\pi^2) - C_{OB} - \sum b(e_i)\pi_i^2)(1 - R(\pi)) + (C_{OB} \]
\[ + \sum b(e_i)\pi_i - C'_{CM})(1 - R(\pi)) + g - R(\pi)\sum b(e_i)(\pi_i^2 - \pi_i) \]  \hspace{1cm} (54)

Note that \( b(\pi^2) - C_{OB} - \sum b(e_i)\pi_i^2 \leq 0 \). Also, by the condition of the claim, the remaining term is also negative. Therefore, we get \( b(\pi^2) - b(\pi) - \sum b(e_i)(\pi_i^2 - \pi_i) < 0 \) under the assumption of \( \delta^*(\pi) = NA \). However,

\[ b(\pi^2) - b(\pi) - \sum b(e_i)(\pi_i^2 - \pi_i) \]  \hspace{1cm} (55)
\[ = b_{OB}(\pi^2) - b(\pi) - b_{OB}(\pi^2) + b_{OB}(\pi) \]  \hspace{1cm} (from \( \delta^*(\pi^2) = OB \)),
\[ = -b(\pi) + b_{OB}(\pi) \geq 0, \]  \hspace{1cm} (56)

which contradicts the assumption. As a result, \( \delta^*(\pi) \) cannot be \( NA \). Also note that \( b_{OB}(\pi) \leq b_{OB}(\pi^2) \leq b_{PM}(\pi^2) = b_{PM}(\pi) \). Therefore, \( PM \) cannot be also optimal, which concludes \( \delta^*(\pi) = OB \).
Proof of Corollary 3:
(a) By Proposition 4(b), $\pi \prec_{st} \pi^2$. Now applying Proposition 4(a) repeatedly to both sides of this inequality yields the result.
(b) Since the states along any sample path is in $\prec_{st}$-increasing order, the result follows directly from Lemma 1.
(c) Note that $R(\pi^k)$ is non-increasing in $k$ by proposition 3(a). Then, the result follows from Lemma 2.
(d) For $k \geq k_1(\pi)$, NA cannot be the optimal action from Lemma 2. Also, for $k \geq k_2(\pi)$, PM is preferable to OB since $C_{OB} + \sum b(e_i)\pi^k_i$ is nondecreasing in $k$ in a $\prec_{st}$-increasing sample path and $C_{PM}^*$ is constant. Hence for $k \geq k^*$, either NA or OB cannot be optimal. For $k_1 \leq k < k^*$, OB is optimal, whereas $k_2 \leq k < k^*$, NA is optimal. For $k < \min\{k_1, k_2\}$, OB or NA is optimal.

Proof of Lemma 5:
We apply the similar induction technique used in [9]. Suppose that $C_{M_0}(e_{m+1}) = C_{CM}$. Also, suppose that $V_0(\pi) = 0$ for $\forall \pi$ for an operating system. $NA_1(\pi) = C_{CM}(1-R(\pi))$ is linear in $\pi$. $OB_n(\pi)$ is hyperplane of $\pi$ and $PM_n(\pi)$ is constant in $\pi$ for $\forall n$. Therefore, $V_1(\pi)$ is piecewise linear concave because minimum of linear functions is piecewise linear concave. Now, suppose that $V_n(\pi)$ is piecewise linear concave such that $V_n(\pi) = \min\{\pi \cdot a_n^T; a_n \in A_n\}$ where $a_n$ is a $1 \times (m+1)$ dimensional column vector. We only need to examine $NA_{n+1}(\pi)$ to show the piecewise linear concavity of $V_{n+1}(\pi)$. The first term of $NA_{n+1}(\pi)$, (that is, $(\tau T + CM_{n-T-1}(e_{m+1}))(1-R(\pi)))$ is linear in $\pi$. The second term of $NA_{n+1}(\pi)$ is,

$$R(\pi)V_n(\pi^2) = R(\pi)\min\{\pi^2 \cdot a_n^T; a_n \in A_n\}$$

$$\quad = R(\pi)\min\left\{\left[\begin{array}{c} (\pi P)_1 \\ (\pi P)_2 \\ \vdots \\ (\pi P)_m \\
\end{array}\right] \cdot a_n^T; a_n \in A_n\right\}$$

$$\quad = \min\{[(\pi P)_1, (\pi P)_2, \ldots, (\pi P)_m, 0] \cdot a_n^T; a_n \in A_n\}$$

$$\quad = \min\{\pi \cdot a_{n+1}^T; a_{n+1} \in A_{n+1}\}$$

Since $R(\pi)V_n(\pi^2)$ is the minimum of hyperplanes, it is piecewise linear concave, which makes $NA_{n+1}(\pi)$ is also piecewise linear concave. Consequently, $V_{n+1}(\pi)$ is piecewise linear concave. And the claim holds for $\forall n$ by induction.

Proof of Theorem 2:
Consider the two states $\pi(\lambda_1)$ and $\pi(\lambda_2)$ between $\pi$ and $\hat{\pi}$ (in $\prec_{st} \hat{\pi}$) where $\pi(\lambda_j) = \lambda_j \pi + (1 - \lambda_j)\hat{\pi}$, for $j = 1, 2$ and $0 \leq \lambda_1 \leq \lambda_2 \leq 1$. Then, from $\sum_{i \geq j} \pi_i \prec_{st} \lambda_1 \sum_{i \geq j} \pi_i + (1 - \lambda_1) \sum_{i \geq j} \hat{\pi}_i \prec_{st} \sum_{i \geq j} \hat{\pi}_i$, we have $\pi \prec_{st} \pi(\lambda_1) \prec_{st} \hat{\pi}$. In a similar way, we can easily show that $\pi(\lambda_1) \prec_{st} \pi(\lambda_2) \prec_{st} \hat{\pi}$. Therefore, $\pi(\lambda)$ is in $\prec_{st}$-increasing in $\lambda$, which implies that $b_{NA}(\pi(\lambda))$ and $b_{OB}(\pi(\lambda))$ is non-decreasing in $\lambda$. But, $b_{PM}(\pi(\lambda))$ is constant. Hence, there exists a control limit $\lambda^*$ such that for any $\lambda > \lambda^*$, $PM$ is optimal. The value of $\lambda^*$ is straightforward from Theorem 1. Next, let’s consider $0 \leq \lambda \leq \lambda^*$. For this region, we already know that $PM$ cannot be optimal from Theorem 1. In Lemma 5, we show that $NA_0(\pi)$ is piecewise linear concave. Thus $b_{NA}(\pi)$ is also piecewise linear concave, but $b_{OB}(\pi)$ is hyperplane. Thus, $\{\pi; b_{NA}(\pi) \geq b_{OB}(\pi)\}$ is a convex set and thus, $\{\lambda; b_{NA}(\pi(\lambda)) \geq b_{OB}(\pi(\lambda)), 0 \leq \lambda \leq \lambda^*\}$ is also a convex set. This concludes the AM4R structure.
**Proof of Corollary 4:**
When $\lambda_{NA} \leq \lambda_{PM} < \lambda_{OB} \leq \lambda_{PM}$, the second $NA$ region of AM4R structure vanishes. So the optimal policy structure results in at most three regions. The optimal policy region for $PM$ is straightforward from the previous discussions.

**REFERENCES**


