Joint optimization of ordering and maintenance with condition monitoring data

As Published: https://doi.org/10.1007/s10479-017-2745-3

Publisher: Springer US

ANOR manuscript No.
(will be inserted by the editor)

Joint Optimization of Ordering and Maintenance with Condition Monitoring Data

Ramin Moghaddass · Şeyda Ertekin

Abstract We study a single-unit deteriorating system under condition monitoring for which collected signals are only stochastically related to the actual level of degradation. Failure replacement is costlier than preventive replacement and there is a delay (lead time) between the initiation of the maintenance setup and the actual maintenance, which is closely related to the process of spare parts inventory and/or maintenance setup activities. We develop a dynamic control policy with a two-dimensional decision space, referred to as a warning-replacement policy, which jointly optimizes the replacement time and replacement setup initiation point (maintenance ordering time) using online condition monitoring data. The optimization criterion is the long-run expected average cost per unit of operation time. We develop the optimal structure of such a dynamic policy using a Partially Observable Semi-Markov Decision Process (POSMDP) and provide some important results with respect to optimality and monotone properties of the optimal policy. We also discuss how to find the optimal values of observation/inspection interval and lead time using historical condition monitoring data. Illustrative numerical examples are provided to show that our joint policy outperforms conventional suboptimal policies commonly used in the literature.

Keywords Real-time Analytics, partially observable semi-Markov decision process, condition monitoring, deteriorating systems

1 Introduction

Mechanical systems often tend to deteriorate over time and with usage due to some underlying degradation processes, random shocks, and the effects of internal and external factors (e.g., load, stress, environmental factors). One of the most important decisions that many companies face is when to turn off mechanical equipment in order to perform preventive maintenance. Considering wind farm maintenance for instance (similar problems occur for oil drilling equipment and electrical feeders, etc.), it is much more cost effective to shut a turbine off before it fails than to repair extensive damage caused by a failure. The goal then becomes one of prediction: if we stop the turbine too early before it would have failed, we lose valuable operating time. If we stop it too late, the turbine may have sustained a catastrophic failure that is expensive to repair. Recent advances in sensor technology and condition monitoring (CM) frameworks have enabled manufacturing systems to monitor the health of operating components continuously without having to suspend the operation of such systems for inspection. For example, like many other types of large mechanical equipment (e.g., oil drilling equipment, electrical feeders), wind turbines are usually equipped with supervisory control and data acquisition (SCADA) sensors that record various measurements of the dynamic environment every few minutes. The information collected through CM, which can be in the form of SCADA data, vibrations measurements, oil analysis, etc., often gives only partial (imperfect) information with regards to the true degradation state.

[1] R. Moghaddass, Department of Industrial Engineering, University of Miami, E-mail: ramin@miami.edu ·
[2] Ş. Ertekin, Orta Dogu Teknik Universitesi & Massachusetts Institute of Technology, E-mail: seyda@ceng.metu.edu.tr

A long-standing problem in operations management and maintenance optimization literature is how to use real-time CM data collected over time for decision-making. The important decision variables we are faced with can be categorized into (i) dynamic decision variables that are correlated with real-time CM data and (ii) static decision variables that are fixed and can be determined offline. The main dynamic decision variables are the replacement time and the warning time (that is the time of the initialization of the replacement setup), and the main static decision variables are the observation/inspection interval (sampling frequency), and the lead time for maintenance setup. It is clear that the above variables are strongly interconnected and should be optimized simultaneously. Although, many research work has been devoted to investigate how to employ real-time information to find the optimal solution of the above variables, very few works regarding the joint optimization of multiple decision variables have been published, particularly for partially observable degrading systems under condition monitoring.

In this paper, we study a condition-monitored deteriorating system with properties that are widely accepted and applicable in the domain of reliability and maintenance. The CM signals (which may only provide partial information with regards to the degradation level) are collected at discrete time points with interval δ, which varies depending on the system and the type of the sensor used for data collection. It is obvious that the smaller the δ, the more CM signals we collect over time, which can potentially lead to better accuracy for system diagnostics and prognostics but higher computational complexity due to the fact the more data needs to be processed. The system is operating until it fails or preventively replaced. Replacement costs include the cost of a new spare plus the additional failure cost (extra labor, damage to the rest of the system, etc.). Since replacement requires some maintenance setup activities, such as gathering tools, materials, labors, ordering spares, labors, the slack time to the generation of work orders, there is a delay (lead-time or maintenance setup time) between the initiation of maintenance setup and the actual start of maintenance. This lead time, denoted by l, which is closely related to the process of spare parts inventory and maintenance setup activities, will influence the downtime of the system. The terms lead time and maintenance setup time are used interchangeably throughout the paper. Spare part ordering (which can be considered as one important step of maintenance setup) with a positive lead time, along with maintenance scheduling, is a crucial decision-making issue in the prognostics and health management (PHM) field. This kind of system is usually expensive and operates singly, and not more than one spare part is kept in stock at the same time (Wang et al, 2015). The existence of the positive lead time between the maintenance initiation point and actual maintenance makes the maintenance setup initiation point a decision variable, which is referred to as the warning time or ordering time in this paper. There is a cost associated with a late warning (costs related to the unavailability of the system or missed profit while the system is waiting for maintenance setup to complete) and early warning (opportunity costs related to maintenance setup finished too early, holding the spare in inventory, depreciation, etc.). In addition, decisions are made at certain decision epochs. The decision interval can be smaller than observation interval. It is a-clear that the maintenance setup initiation time can affect the replacement time and thus these variable are highly correlated and represent two distinct and dependent actions.

We first develop a dynamic, joint warning-replacement policy using partially observable semi-Markov decision process, which has a rich mathematical framework. The proposed decision policy is a two-dimensional control policy that employ the history of collected CM data to determine two types of actions at each decision epoch, (i) whether or not to initiate maintenance setup and (ii) whether or not to terminate the operation and replace the degraded system with a new one (or wait until the maintenance setup is complete). We then describe simple approaches that can be used to find the best set of observation interval and lead time among available alternatives. The results obtained from this paper can also be used to determine the effectiveness of a condition monitoring system with respect to a condition-based maintenance with direct inspection data or no monitoring data at all. This work is motivated by the increased interest of the wind turbine industry to use the SCADA (Supervisory Control And Data Acquisition) measurements, which may give only partial information with respect to the degradation process of a wind turbine, to issue better timed warnings (maintenance setup initiation points) and to determine the optimal time to terminate the operation for maintenance. In the context of inventory management, our model can be considered as a dynamic joint ordering-replacement policy in a single-unit inventory system, considering that there is a delay between the placement of the order and delivery of the spare (i.e., positive ordering lead time). When warning corresponds to ordering, early warning cost is related to inventory holding costs and late warning cost is related to shortage costs. In this paper, the terms ordering and warning are used interchangeably. We should point out that the results of this paper are applicable only if the cost of early warning and/or late warning is non-zero, that is completing maintenance setup early or late (whether it relates to spare part inventory or other maintenance setup activities) incur a considerable cost. Fig. 1 shows a simulated CM signal and two decision points of warning and replacement. In this example, warning and replacement points are both before the actual failure point. In addition, the warning point plus the lead time is right before the replacement point and therefore the system does not have to wait until the setup is finished. It is obvious that warning point and replacement point are highly correlated. The goal of this paper is to develop a method that can use this type of online CM signal to make decisions that help avoid too late and too early warnings (start of maintenance setup or o

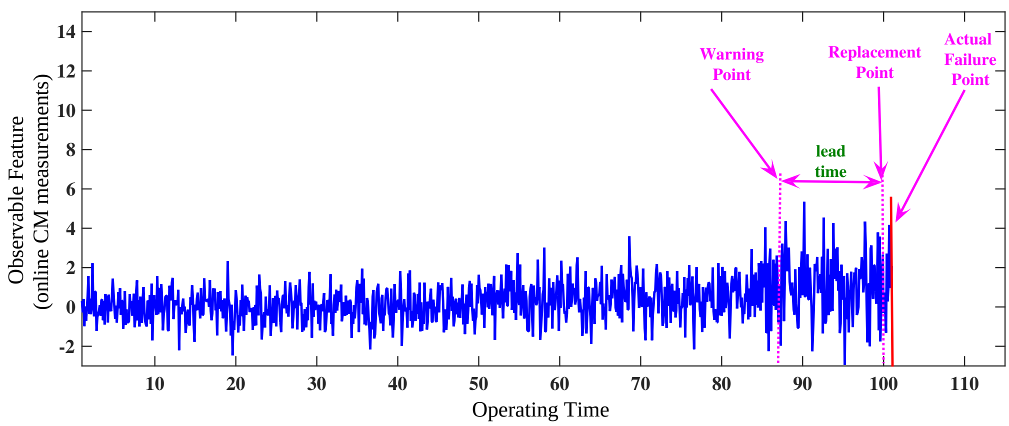

Fig. 1: Sample online CM data and decision points.

The remainder of the paper is organized as follows: In §2, we review some of the relevant literature on CBM optimization, joint inventory-maintenance optimization models, and partially observable degrading systems, and summarize our contributions made in this paper. The preliminaries and the problem setting are discussed in §3. In §4, we present our dynamic control model for the joint optimization of warning time and replacement time. In §5, the structure of the optimal policy is analyzed. In §6, how to find static decision variables from multiple alternatives is discussed. We also describe important special cases with regards to the availability of condition monitoring data. A numerical example is given in §7 to illustrate our model and its possible benefits. Finally, concluding remarks and future extensions are outlined in §8.

2 Literature Review

Optimization of parameters of inventory and maintenance of deteriorating systems is a well-studied problem in the literature. One of the earliest works on joint warning-replacement policy is the series of papers by Armstrong and Atkins (1996, 1998) where they examined an age-based replacement and ordering decisions, for a single-unit system with a lifetime following a known probability density function (PDF). The long-run average costs were minimized by searching the optimal combination of replacement time and ordering time. With respect to optimization parameters, most of the literature have focused on equipment replacement time and spare part inventory ordering time. A large body of literature has been devoted to either optimizing one variable at a time or optimizing sequentially all decision variables. Examples of models that focused only on replacement time can be found in Rosenfield (1976); Ivy and Pollock (2005); Maillart (2006); Chen and Wu (2007); Kurt and Kharoufeh (2010) and examples of models that focused only on ordering time can be found in Louit et al (2011); Godoy et al (2013). For a comprehensive review of the available literature on this topic, interested readers may refer to the work of Van Horenbeek et al (2013).

Most of the conventional models in the literature, which falls in the category of age-based maintenance (e.g., Armstrong and Atkins, 1996) and block-based or periodic preventive maintenance (e.g., Brezavscek and Hudoklin, 2003), has the limitation that the maintenance actions are set for all components regardless of the actual degradation pattern and failure history each may experience while in operation. This shortcoming can potentially lead to too many unnecessary maintenance actions, unexpected failures, and too early/too late orders (warnings). That is what motivates researchers in recent years to combine CBM and optimization of inventory and maintenance parameters in order to cover more realistic situations. The concept of CBM has been applied extensively in the domain of reliability and maintenance (e.g., Jardine et al, 2006; Zhang et al, 2006; Peng et al, 2010; Ahmad and Kamaruddin, 2012). Although using CM data for decision-making seems to be the case in many application settings, very few have considered using online CM data to jointly optimize the maintenance and inventory policies.

An example of CM sensor-driven prognostic models for joint inventory-maintenance decision-making is the work of Elwany and Gebraeel (2008) under which the optimal replacement time is first optimized followed by the optimal ordering time. In other words, once the optimal replacement time has been computed, it is used to decide when to order a spare part. Their work is the extension of the well-known work of Armstrong and Atkins (1996), with the difference that sensor measurements (which directly reflect the degradation level) are used to update the remaining lifetime distribution over time. It has been shown by many researchers (e.g., in Kabir and Farrash, 1996) that sequential optimization does not guarantee global optimality. Our work is different from Elwany and Gebraeel (2008), as (i) we do not assume that CM data give perfect information with regards to the degradation process, (ii) we jointly optimize ordering time and replacement time using dynamic programming, and (iii) we develop the structure of the optimal policy. In addition, our model is more general with respect to the degradation process and we do not assume a known threshold for failure in terms of observed signals.

Panagiotidou (2014) studied the joint maintenance and ordering problem for more than one identical items with Markovian degradation, which are under periodic inspection that could reveal the true state of the system perfectly. At each inspection point, the system is restored to an as-good-as-new condition. Two ordering policies namely a periodic review policy (a multiple of the inspection interval) and a continuous review policy (order when available spares reach a predefined limit) were investigated. They did not investigate the structure of the joint optimal policy. Their model is useful only when the lead-time is smaller than the inspection interval. Wang et al (2008a) proposed a cost-effective condition-based order-replacement policy to jointly optimize the inspection interval, the ordering threshold, and the preventive replacement threshold for a single unit system with monotonically increasing but completely observable degradation process with a known structure for the degradation process. They suggested a fixed threshold policy in which the ordering time and replacement time were determined based on the observed level of degradation. They extended their model in Wang et al (2008b) to a system with multiple identical units and used Monte Carlo (MC) simulation to evaluate the cost rate. Finally, in Wang et al (2009), they presented a condition-based replacement policy with periodical inspections for deteriorating systems with a number of identical units. They combined the condition-based replacement policy with periodical inspections and the (S, s) type inventory where preventive replacement threshold was defined in terms of the deterioration levels of units and ordering was defined in terms of the level of inventory. Our work is different from all of the above as we develop a dynamic decision process that simultaneously optimizes both the warning time and replacement time where only partial information is available with regards to the degradation process. Also, in our model we investigate the structure of the optimal policy as opposed to having a pre-determined structure for the optimal policy.

With regards to the observation process, most of the literature on CBM has assumed a completely observable systems in which the CM signals obtained at predetermined monitoring/inspection points either directly reflect the level of degradation or provide perfect information with respect to the degradation state (e.g., Lam and Yeh, 1994; Yeh, 1997). For example in the well-known work of Banjevic et al (2001), the authors have assumed that the covariate measurements, which are observable at certain discrete points over time reflect the level of degradation directly. Many extensions of this work and this assumption have been published including the recent work of Wu and Ryan (2014) and Qian and Wu (2014). Our work is different from this group of literature, as we consider a more realistic case under which condition monitoring data provide only partial information with regards to the degradation process and are only stochastically related to the underlying system state. For a review of CBM models on partially observable degrading systems, interested readers may refer to Jardine et al (2006). We should remind here that the degrading system with perfect CM information can be considered as a particular case of partially observable degrading systems.

The models developed in this paper contributes both to the advancement of reliability analysis and condition-based maintenance for degrading systems and the advancement of partially observed Markov and semi-Markov decision processed with dependent decision variables. The contribution made in this paper is three-fold. First, we develop a dynamic control policy using a partially observable semi-Markov decision process that provides a series of optimal actions for replacement and warning (ordering) for a single-unit condition-monitored device. Unlike most of the reported works in this area, we investigate the structure of the optimal policy. Second, we assume a more generic assumption with regards to condition monitoring data, that is, condition monitoring data give only partial information with regards to the actual degradation process. The history of collected CM data is employed at each decision epoch to dynamically update the health status of the system. In terms of the degradation process, we consider a very general structure that is widely used and accepted in the literature. To our knowledge, this is the first time that a dynamic decision model is introduced that can employ online CM data for joint optimization of warning time and replacement time. Third, we describe a simple approach to investigate optimal monitoring/inspection time and optimal lead time from several alternatives. In addition to the above, our model can be used to analyze the two extreme cases of a) perfect CM obtained by inspection, and b) no CM at all. Such analyses can be employed to determine whether or not monitoring the health of the system over time is beneficial. We also discuss special cases, where we have to deal with missing data, outliers, large monitoring intervals, and fully observable degradation processes. We develop a partially observable semi-Markov decision process and present some structural properties of the optimal policy, such as monotonicity and control-limit form. We should point out that unlike most of available literature (e.g., Armstrong and Atkins, 1996; Elwany and Gebraeel, 2008), we do not force the ordering time plus lead time to be smaller than the replacement time, instead we let the dynamic model determine the decision variables.

3 Preliminaries & Problem Settings

In this section, the main characteristics of the degradation process, the observation process, cost elements, decision variables, and the assumptions made throughout the paper are described.

3.1 Evolution of the Degradation Process

The deteriorating system considered in this paper is associated with a single unit that operates until it fails or preventively replaced. The degradation process considered here has a very general form that is widely accepted and frequently used for degradation modeling (for example in Banjevic et al, 2001; Banjevic and Jardine, 2006). Let $Z = \{Z(t) : t \in \mathbb{R}^+\}$ be a continuous-time stochastic process with discrete, finite state space $\mathscr{E} = \{0, 1, \cdots, N\}$ that is related to the overall state of health of the system and can influence the total hazard rate. States are ordered to reflect the relative degrees of deterioration of the system. The process $Z(t)$ can also be a diagnostic covariate that reflects directly or indirectly different levels of degradation. For multidimensional values of Z, the total state space can be defined based on all possible combinations of individual states (Louit et al, 2011). Examples of multi-state stochastic processes that reflect degradation are the 5-state (sharp, normal wear, micro fracture, macro wear, and breakage) degradation process for a friction drilling device used in Hsu and Shu (2010) and the 4-state (baseline, contamination 1, contamination 2, failure) degradation process defined for hydraulic pump health monitoring in Teng et al (2011). A very common approach to deal with process $Z(t)$ is to assume that it changes only at discrete points of time denoted by $\{0, \delta, 2\delta, ...\}$. Then the right continuous jump process $Z_k, k \in \mathbb{N}_0 = \{0, 1, 2, ...\}$ can be used as the approximation of the stochastic process $Z(t), t \in \mathbb{R}^+$, where Z_k is the value of the stochastic process Z at time $k\delta$. The evolution of the stochastic process Z_k in discrete domain is governed by a stochastically increasing matrix $\mathbf{P}(k) = [p_{ij}(k)], i, j \in \mathscr{E}, k \in \mathbb{N}_0$, where $p_{ij}(k)$ is the conditional probability of transition from state i to state j at time $k\delta$, given that the system has survived until time $k\delta$. This probability can be constant or covariate-dependent (e.g., time-dependent, state dependent). Now, let us also define $\lambda(t, Z(t))$ as the conditional hazard rate (failure rate) at time t, given $Z(t)$. It is very common in the literature of hazard modeling that this hazard function is represented in terms of a baseline hazard function $\lambda_0(t)$ and a link function that depends on the state of the stochastic covariate process Z (e.g., Makis and Jardine, 1992). Then the following holds true for the joint conditional distribution of the lifetime and the state of the process in the discrete domain, given that the system has survived up to time $k\delta$:

$$\Pr(\zeta > (k+1)\delta, Z_{k+1} = j | \zeta > k\delta, Z_k = i) = \exp\left\{-\int_{k\delta}^{(k+1)\delta} \lambda(t, Z(t))dt\right\} p_{ij}(k+1),$$

where ζ is the random variable denoting the life of the system. Now the conditional reliability function $R(t|k\delta, i) = \Pr(\zeta > t|\zeta > k\delta, Z_k = i)$ given the state of the system Z_k can be defined as

$$R(t|k\delta, i) = \begin{cases} \exp\left(-\int\limits_{k\delta}^{t} \lambda(\tau, i)\, d\tau\right), & t \leq (k+1)\delta \\ \exp\left(-\int\limits_{k\delta}^{(k+1)\delta} \lambda(\tau, i)\, d\tau\right) \sum\limits_{j\in\mathscr{E}} p_{ij}(k+1) R(t|k\delta+\delta, j), & t > (k+1)\delta \end{cases}. \tag{1}$$

The results presented in this paper are based on the assumptions given in Ohnishi et al (1994) that (i) the state of the process Z is stochastically increasing in time and as the degradation level becomes higher, it is more likely to make a transition to a higher degradation level (i.e., $\mathbf{P}$ is totally positive of order 2 or TP2 in short), and (ii) the hazard function $\lambda(t, Z(t))$ is a nondecreasing function in t and the state of the process $Z(t)$, that is, the system is more likely to fail at a higher level of Z and t. These assumptions are commonly used in the literature for non-repairable systems under gradual degradation. We should point out that our model can be applied to any deteriorating system where it is possible to represent hazard rate and probability of transition in terms of a combination of time, degradation level, and other measurable covariates.

3.2 The Observation Process (Condition Monitoring Process)

As stated earlier, the process Z is not directly observable, and only partial information can be inferred from it with respect to the state of the system. The observation process $Y_k, k \in \mathbb{N}_0$, which takes values from the finite set $\mathcal{Y} = \{1, 2, ..., M\}$ and is sampled at discrete points $\{0, \delta, 2\delta, ...\}$, is stochastically dependent on the true state of the system. Examples of the observation signal used to monitor degradation are in Dong and He (2007) where vibration signals were fused based on the importance of sensors data to monitor the degradation of hydraulic pumps and in Sun et al (2012) where a single health index (HI) is inferred from a set of sensor signals to characterize the hidden health state of the system for turbofan engine degradation. While the system is in state i, the probability that output j is observed is $b_i(j)$, which belongs to a known distribution $\mathbf{B} = [b_i(j)], i \in \mathscr{E}, j \in \mathcal{Y}$, referred to as the state-observation matrix (information matrix). This form of observation matrix has been considered many times in the literature (see for instance Ghasemi et al, 2010; Kim and Makis, 2013). Let us define θ_k^i as the conditional probability of being in state i given the series of observation signals $y_1, y_2, ..., y_k$. Then we have

$$\theta_k^i = \Pr(Z_k = i|Y_1 = y_1, Y_2 = y_2, ..., Y_k = y_k) = \frac{\Pr(Z_k = i, Y_k = y_k|Y_1 = y_1, Y_2 = y_2, ..., Y_{k-1} = y_{k-1})}{\Pr(Y_k = y_k|Y_1 = y_1, Y_2 = y_2, ..., Y_{k-1} = y_{k-1})} \tag{2}$$

$$= \frac{\sum\limits_{j\in\mathcal{E}} \Pr(Z_{k-1} = j|Y_1 = y_1, Y_2 = y_2, ..., Y_{k-1} = y_{k-1}) \times p_{ji}(k) \times b_i(y_k)}{\sum\limits_{i\in\mathcal{E}} \sum\limits_{j\in\mathcal{E}} \Pr(Z_{k-1} = j|Y_1 = y_1, Y_2 = y_2, ..., Y_{k-1} = y_{k-1}) \times p_{ji}(k) \times b_i(y_k)},$$

with the initial state condition $\boldsymbol{\theta}_0 = [\theta_0^1, \cdots, \theta_0^N]$, where $\theta_0^i = \Pr(Z_0 = i)$, $\forall i \in \mathscr{E}$. Now, given that the series $\{y_1, y_2, ..., y_k\}$ is fully observed, the conditional reliability in terms of $\boldsymbol{\theta}_k = [\theta_k^1, \cdots, \theta_k^N]$ can be computed as

$$R(t|k\delta, \boldsymbol{\theta}_k) = \Pr(\zeta > t|\zeta > k\delta, Y_1 = y_1, Y_2 = y_2, ..., Y_k = y_k) = \sum_{i\in\mathcal{E}} \theta_k^i R(t|k\delta, i). \tag{3}$$

Based on the results given in Ghasemi et al (2007), we also assume that $R(t|k\delta, \boldsymbol{\theta}_k)$ is non-increasing in k and $\boldsymbol{\theta}$, that is larger age and larger distribution of the deterioration level in the sense of stochastic ordering are likely to yield higher (or equal) probability of failure. Similar to Ohnishi et al (1994) and Ghasemi et al (2007), it is also assumed that $\mathbf{B}$ is TP2, which means that higher degradation level yields higher degradation signal stochastically.

3.3 Cost Elements

The replacement of the device (either a failure replacement or a preventive replacement) costs c_r. There is an additional cost for a failure replacement at state i denoted by c_f^i, which covers all relevant charges including extra labour, repair of damage to the rest of the system, cost of downtime, etc. The failure cost vector $\boldsymbol{c}_f = [c_f^1, \cdots, c_f^N]$ is assumed to be nondecreasing ($c_f^1 \leq c_f^2 \leq \cdots \leq c_f^N$), that is, the failure cost is larger for a higher degradation level. The system is monitored at discrete time points referred to as observation epochs $(\delta, 2\delta, \ldots)$, where δ is the observation interval. For simplicity, we assume that decision-making interval is equal to the observation interval, however we later explain how to relax this assumption. There is a delay (lead time) between the start of the maintenance setup (which will be referred hereafter as the warning point) and the completion of the replacement setup, denoted by l. This replacement setup can be related to spare ordering or maintenance initiation setup. If the system is unavailable because the replacement setup is not ready (e.g., when the spare order is not delivered), a shortage cost of c_s, per unit time representing missing profit or downtime cost is incurred. On the other hands, if the maintenance setup is ready (or if the spare order is ready) before the system fails, a holding cost of c_h per unit time is incurred. This cost can represent the opportunity cost of capital, depreciation of the spare, or the cost of having maintenance setup earlier than needed. It is assumed that the fixed cost of ordering or maintenance setup (per unit) when the lead time is l is c_l. For notational convenience, we define function $\mathscr{C}_l(t, d)$ as the cost of ordering at time d and termination of the operation for replacement at time t as

$$\mathscr{C}_l(t, d) = c_s(d + l - t)\mathbb{1}_{\{t \leq d+l\}} + c_h(t - d - l)\mathbb{1}_{\{t \geq d+l\}}. \tag{4}$$

Based on the above definition, if the warning process starts exactly l units before the failure (that is when $t - d = l$), then $\mathscr{C}_l(t, d)$ becomes zero. In summary, the following reasonable assumptions are made in this cost function: (i) if the warning is issued l units before the failure, then the cost of the warning process is zero, (ii) both late warning and early warning are costly, that is, it takes c_s and c_h for each units of delay warning and early warning, respectively, (iii) late warning is more costly than early warning ($c_s > c_h$), and (iv) the cost of warning at failure is greater than the cost of warning at time zero, that is, early warnings are preferred over late warnings. We used this form of cost function due to its mathematical simplicity, high interpretability, and its common use in the literature. Examples of cost function in the same form as ours in the context of joint ordering and replacement can be found in Elwany and Gebraeel (2008); Panagiotidou (2014).

3.4 Decision Variables

The model developed in this paper can be used to optimize two types of decision variables. The first type, which is the main focus of this paper, includes dynamic decision variables, namely, warning time (which can be related to either the spare part ordering or maintenance initiation setup) and replacement time, which are determined through the series of actions chosen at decision epochs $\delta, 2\delta, \ldots.$, at which samples of the observation signal are taken. The second group deals with static decision variables including the observation interval and the lead time. In Phase I (described in §4), the structure of the dynamic control model to find the series of actions associated with dynamic decision variables is described. Then in Phase II (described in §6), a simple approach to find the best set of static decision variables among available alternatives is described.

4 The Semi-Markov Control Model for the Development of Joint Warning-Replacement Policy

The proposed dynamic control model has 3 elements: state space, maintenance actions, and a control policy.

4.1 State Space

The measurable state of the system at the kth decision point, denoted by $\boldsymbol{\pi}_k$, is defined as a 4-dimensional vector that can fully represent all information needed for decision making. The element of $\boldsymbol{\pi}_k$ are (1) the age of the system, (2) the belief state $\boldsymbol{\theta}_k$, (3) the status of the warning generation process (o_k), and (4) the working status of the system (υ_k). The first element of $\boldsymbol{\pi}_k$ is the age of the process, which can influence the decision process as well as the evolution of covariate Z and failure rate λ. In a regular case, the age of the system at the kth decision epoch is $k\delta$, but to allow more flexibility, we can let the age be any positive value. As the process Z is not directly observable over time, we cannot not include it in the state space of the decision process. Instead, we employ the belief state (information vector). The set of all possible belief states is called the belief space Θ as defined below

$$\boldsymbol{\theta}_k \in \Theta = \left\{ \{a_1, \cdots, a_N\}\,;\, a_i \in \mathbb{R}^+, 0 \le a_i \le 1, \forall\, i \in \mathscr{E}, \sum_{i=1}^{N} a_i = 1 \right\}, \forall k = 1, 2, \dots .$$

The second element of $\boldsymbol{\pi}_k$ is the belief state $\boldsymbol{\theta}_k = [\theta_k^1, \theta_k^2, \dots, \theta_k^N]$ as the conditional probability distribution of the degradation level at time $k\delta$ (see Eq. (2)). We should remind that this so-called *belief state* is fully observable over time (measurable), since it is a function of the known initial distribution and the observation process. We also define $o_k (o_k \in \{\mathbb{N} \cup \infty\})$ as an element of the state indicating the time point at which the warning is issued. By definition, $o_k = o$ means that the warning was issued at the oth decision point. If the warning has not been issued yet, then we let $o_k = \infty$. The last element of the state of the process is the overall status of the system indicating whether the system is available or failed. We define υ_k ($\Upsilon_k \in \{1, 0\}$), where 0 means failure and 1 means working. For example, state $(20, (0.2, 0.7, 0.1), 10, 1)$ means the age of the system is 20 units, the probability of being at states 1-3 are 0.2, 0.7, and 0.1, respectively, the warning has been issued at the 10 decision point, and the system is still operating. Based on the above elements of the state, the fully observable state of the decision process at time $k\delta$ denoted by vector $\boldsymbol{\pi}_k$ has $N + 3$ elements from which N are for the belief state and 3 are for age, warning time, and working status. Therefore, the state space $\mathcal{S}$ is

$$(k, \boldsymbol{\theta}, o, \upsilon) \in \mathcal{S} = \mathbb{N}_0 \times \Theta \times \{\mathbb{N}_0 \cup \{\infty\}\} \times \{1, 0\}\,, \text{where } \mathbb{N}_0 = \{0, 1, 2, \dots\}.$$

As the conditional state probability $\boldsymbol{\theta}$ is fully defined through the observation process, which is finite by definition, the space $\mathcal{S}$ (if only feasible points are considered) can be converted to a countably finite set with possibly many states. In Section §5.5, we describe how we can manually discretize the interval [0,1] for the belief state, so that $\mathcal{S}$ becomes a much smaller finitely countable discrete space. It will be shown later that $\boldsymbol{\pi}_k$ is a sufficient statistic for the dynamic control of the replacement and warning processes.

4.2 The Set of Possible Control Actions

The control set or action space includes four 2-dimensional vectors $\mathcal{U} = \{(1, 1), (0, 1), (1, 0), (0, 0)\}$. For each $u_k = \{u_k^1, u_k^2\} \in \mathcal{U}$, u_k^1 is the action made at the kth decision point corresponding to the warning generation process (0 means issue warning immediately and 1 means do nothing, that is wait until the next decision point) and u_k^2 is the action corresponding to the replacement process (0 means replace immediately and 1 means wait until the next decision point). Some of these actions are not admissible/feasible for certain states. The list of admissible actions under different state conditions is given below. For example, the action $(1, 0)$ is not meaningful when the warning is not issued yet (that is, warning time should not exceed the replacement time). At a decision point, if the system is found to be failed, then it is replaced immediately and the warning is issued if not issued yet.

$$\text{Admissible actions for state } (k, \boldsymbol{\theta}, o, \upsilon): \begin{cases} \{(1,1),(0,1),(0,0)\}, & \text{if } \{o=\infty, \upsilon=1\} \\ \{(1,1),(1,0)\}, & \text{if } \{o\neq\infty, \upsilon=1\} \\ (0,0), & \text{if } \{o=\infty, \upsilon=0\} \\ (1,0), & \text{if } \{o\neq\infty, \upsilon=0\} \end{cases}, \forall\,(k, \boldsymbol{\theta}, o, \upsilon) \in \mathcal{S}.$$

4.3 Control Policy

A control policy (or strategy) represents decision rules that specify how controls (actions) are applied at each decision epoch. In other words, it is a function from decision space $\mathcal{S}$ to action space $\mathcal{U}$ that transforms each state in the state space to an action in the associated admissible actions for that state. Since the control policy is 2-dimensional, we use $\gamma = \{\gamma_1, \gamma_2\}$ as the decision rule that indicates the ordering action $\gamma_1(\boldsymbol{\pi})$ and replacement action $\gamma_2(\boldsymbol{\pi})$ when the current state is $\boldsymbol{\pi}$. The data available for decision makers at the kth decision epoch prior to action u_k being applied is $\{y_1, (u_1^1, u_1^2), y_2, (u_2^1, u_2^2), \cdots, y_k\}$ from which $\boldsymbol{\pi}_k$ is fully obtainable. We will show that the quantity $\boldsymbol{\pi}_k$ represents sufficient information for decision-making at the kth decision point.

4.4 The relationship between State Space, Control Actions, and Control Policy

The evolution of the decision process is governed by $\boldsymbol{\pi}_k$, $\boldsymbol{u}_k$, and γ as described below. At the kth decision point, the state of the system $\boldsymbol{\pi}_k$ is fully known. Then the policy γ determines the 2-dimensional action $\boldsymbol{u}_k$ based on $\boldsymbol{\pi}_k$. Then $\boldsymbol{\pi}_k$ and $\boldsymbol{u}_k$ together determine the so-called per-stage cost denoted by $C(\boldsymbol{\pi}_k, \boldsymbol{u}_k)$. This cost includes instantaneous cost plus all costs incurring before the next decision point including the cost of replacement, failure, and warning. We will decompose this cost function in terms of $\boldsymbol{\pi}$ and $\boldsymbol{u}$ in §5.1. The system will then move to a new state $\boldsymbol{\pi}_{k+1}$ with a probability that also depends on $\boldsymbol{\pi}_k$ and $\boldsymbol{u}_k$. This probability, which is defined through a set of transition laws, specifies the evolution of the state space given the previous state and control as a transition probability with the Markovian property as

$$\Pr(\boldsymbol{\pi}_{k+1}|\boldsymbol{\pi}_0, \cdots, \boldsymbol{\pi}_k, \boldsymbol{u}_k) = \Pr(\boldsymbol{\pi}_{k+1}|\boldsymbol{\pi}_k, \boldsymbol{u}_k).$$

By definition, if the action *replace immediately* is chosen at a decision point, the state at the next decision point is always $(0, \boldsymbol{\theta}_0, \infty, 1)$ with probability 1, regardless of the current state, that is, the system goes back to an *as-good-as-new* condition. For other cases where the action do nothing is chosen with respect to replacement, we can use the Baye's theorem to find the posterior distribution of the belief state $\hat{\boldsymbol{\theta}}_{k+1}(m) = [\hat{\theta}_{k+1}^1(m), \cdots, \hat{\theta}_{k+1}^N(m)]$ conditional on observation $y_{k+1} = m$ and the fact that the action at the kth decision epoch is no replacement. This leads to

$$\hat{\theta}_{k+1}^j(m) = \Pr(Z_{k+1} = j|\boldsymbol{\theta}_k, y_{k+1} = m) = \frac{\sum_{i=1}^{N} \theta_k^i \times p_{ij}(k+1) \times b_j(m)}{\sum_{i=1}^{N}\sum_{j=1}^{N} \theta_k^i \times p_{ij}(k+1) \times b_j(m)},\ \forall m \in \mathcal{Y}, \forall j \in \mathscr{E}. \tag{5}$$

The posterior $\hat{o}_{k+1}$, which shows the status of the system at the $(k+1)$th decision epoch in terms of the warning generation process can be determined given o_k and the warning control action u_k^1, as

$$\hat{o}_{k+1}(u_k^1) = \begin{cases} \infty & \text{if } o_k = \infty, u_k^1 = 1 \\ k & \text{if } o_k = \infty, u_k^1 = 0 \\ d & \text{if } o_k = d < k \end{cases}. \tag{6}$$

Equation (6) states that for the cases that warning has not been issued yet ($o_k = \infty$), if the warning control at time k is *do-nothing* ($u_k^1 = 1$), the next state of warning remains ∞. However, if the warning control is *immediate warning* ($u_k^1 = 0$), then the next state of warning becomes k, meaning that the warning is issued at the kth decision epoch. For the cases that warning is already issued at time d ($o_k = d, d < k$), the next warning state remains d. If the replacement control is to replace-immediately, then the system goes back to state 0 and the warning state at the next decision point becomes ∞. Note that the generation of the observation process at the $(k+1)$th point is also affected by the current state of the system and the chosen control for replacement as follows

$$\Pr(y_{k+1} = m|\boldsymbol{\theta}_k, u_k^2) = \begin{cases} \sum_{i=1}^{N}\sum_{j=1}^{N} \theta_k^i \times p_{ij}(k+1) \times b_j(m), & \text{if } u_k^2 = 1 \\ \sum_{j=1}^{N} \theta_0^j \times b_j(m), & \text{if } u_k^2 = 0 \end{cases} \quad \forall m \in \mathcal{Y}. \tag{7}$$

Note that the choice of warning control has no impact on the next observation signal. It can now be summarized that (i) if the action *no replacement* is selected for state $\boldsymbol{\pi}_k = (k, \boldsymbol{\theta}_k, o_k, 1)$, the system moves with probability $R(k\delta + \delta|k\delta, \boldsymbol{\theta}_k) \Pr(y_{k+1} = m|\boldsymbol{\theta}_k, 1)$ to state $(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), \hat{o}_{k+1}, 1)$, $\forall m \in \mathcal{Y}$, and moves with probability $(1 - R(k\delta + \delta|k\delta, \boldsymbol{\theta}_k))$ to state $(0, \boldsymbol{\theta}_0, \infty, 1)$, and (ii) if the action replace immediately is chosen, the system moves to state $(0, \boldsymbol{\theta}_0, \infty, 1)$ with probability 1. Because the interval length between two decision points are not necessarily fixed, we need to define the sojourn time distribution an element of a semi-Markov decision process. This distribution depends only on the replacement control since the choice of control for the warning process cannot influence the sojourn time distribution. By definition, if option *replace immediately* is selected, the system is brought back to the initial condition and therefore the sojourn operating time is theoretically zero. In all other cases, the expected value of the sojourn time depends on the conditional reliability function as follows

$$\bar{\tau}(\delta|k, \boldsymbol{\theta}_k) = \int_{k\delta}^{k\delta+\delta} \bar{R}(\mathrm{d}x|k\delta, \boldsymbol{\theta}_k) + \delta\, R(k\delta + \delta|k\delta, \boldsymbol{\theta}_k) = \int_{k\delta}^{k\delta+\delta} R(x|k\delta, \boldsymbol{\theta}_k)\mathrm{d}x. \tag{8}$$

The graphical model of the described semi-Markov control and its elements are shown in Figure 2.

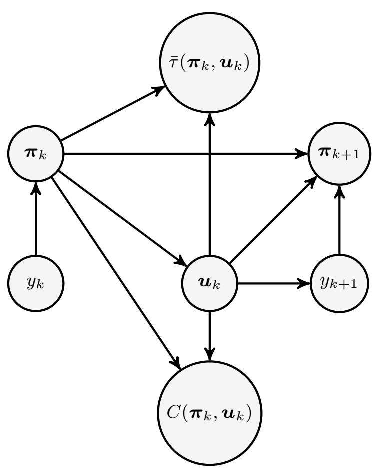

Fig. 2: Graphical model of the associated semi-Markov control and the evolution of the decision process

4.5 Decision-Making Cost Criterion

The objective of this paper is to develop an optimal policy that can minimize the expected long-run average cost per unit time. According to the renewal theory, for an infinite time span, minimizing the expected long-run average cost is equivalent to minimizing the expected cost rate of a maintenance cycle, which is the time interval between two successive replacements (Yeh, 1997). Therefore, the cost function can be defined as the expected cost of one cycle divided by the expected cycle length as defined below.

Definition 1 Let $T_{1,\gamma}$ and $T_{2,\gamma}$ be the warning time and the replacement time associated with policy γ, where $T_{1,\gamma} \leq T_{2,\gamma}$. The expected average cost per unit time of this policy is

$$g(\gamma) = \frac{c_r + \sum_i c_f^i \times \mathbb{E}\Big[\Pr(\zeta \leq T_{2,\gamma}, Z(\zeta) = i)\Big] + \mathbb{E}\left[\mathscr{C}_l(\min(\zeta, T_{2,\gamma}), \min(\zeta, T_{1,\gamma}))\right]}{\mathbb{E}[\min(\zeta, T_{2,\gamma})]}, \tag{9}$$

where $\mathbb{E}$ is the conditional expectation operation, $\mathbb{E}[\Pr(\zeta \leq T_{2,\gamma}, Z(\zeta) = i)]$ is the expected probability of a failure replacement at state i within a replacement cycle, $\min(\zeta, T_{2,\gamma})$ is the effective replacement time, $\min(\zeta, T_{1,\gamma})$ is the effective warning time, $\mathbb{E}\left[\mathscr{C}_l(\min(\zeta, T_{2,\gamma}), \min(\zeta, T_{1,\gamma}))\right]$ is the expected cost of the warning process, and $\mathbb{E}[\min(\zeta, T_{2,\gamma})]$ is the expected cycle duration. The expected probability of a failure replacement in a cycle can be calculated as $\sum_{i=1}^{N} \mathbb{E}\Big(\Pr(\zeta \leq T_{2,\gamma}, Z(\zeta) = i)\Big)$.

The objective of the decision policy is to find the optimal policy γ^* such that $g^* := \inf_\gamma g(\gamma) = g(\gamma^*)$. The optimal average cost function $g^* := \inf_{\gamma \in \Gamma} g(\gamma)$ and a policy γ^* is said to be average cost optimal if $g^* = g(\gamma^*)$ (i.e., $\gamma^* = \underset{\gamma}{argmin}\, g(\gamma)$).

Remark 1 If an optimal policy exists, then its average cost is bounded by

$$\frac{c_r}{\mathbb{E}(\zeta)} \leq g^* \leq \min\left\{\frac{c_r + \sum_i c_f^i \mathbb{E}\left(\Pr(\zeta \leq T_{2,\gamma}, Z(\zeta) = i)\right) + c_s \times l}{\mathbb{E}(\zeta)}, \frac{c_r + \sum_i c_f^i \mathbb{E}\left(\Pr(Z(\zeta) = i)\right) + \mathbb{E}\left(\mathscr{C}(\zeta, 0)\right)}{\mathbb{E}(\zeta)}\right\},$$

where $\mathbb{E}(\zeta)$ is the expected age of the system, and $\mathbb{E}\left(\mathscr{C}(\zeta, 0)\right)$ is the expected cost of the warning generation process if the warning is issued at time 0 and replacement is done at the failure point.

Proof The proof is given in the Appendix.

We will use the upper bound given above as an initial point to find the optimal policy. Decision makers can take the difference between the lower bound and the upper bound as the maximum possible improvement in the average cost function after applying a joint warning-replacement policy. The difference does not necessarily represents the actual improvement, it is just an indicator of the maximum potential of the joint policy in terms of reducing the average cost reduction. If the actual improvement (by applying the proposed optimal policy) equals this theoretical improvement, then we have reached a 100% perfect or an ideal policy. This is extremely unlikely due to the stochastic nature of the degradation and observation processes. This remarks also shows that if the costs of failure and ordering are zero, then the optimal policy is do nothing as the two bounds become the same.

5 Structure of the Optimal Policy

In the previous section, we developed the structure of the semi-Markov model associated with the decision process. In this section, we investigate the structural properties of the optimal policy and develop a simple control-based policy that can be used for the joint determination of warning time and replacement time.

5.1 Structure of the Dynamic Equations

As noted earlier, a warning-replacement policy is a 2-D mapping function from each state $\boldsymbol{\pi} \in \mathcal{S}$ to an action $\boldsymbol{u} \in \mathcal{U}$. Let $V_\gamma(\boldsymbol{\pi})$ be the value function (or the expected relative cost) associated with policy γ in an infinite horizon, given that the decision process starts from $\boldsymbol{\pi} \in \mathcal{S}$. Whenever the system is still operating, the optimality equation at $\boldsymbol{\pi}_k$ becomes

$$V_\gamma(\boldsymbol{\pi}_k) = \min_{\boldsymbol{u}=[u^1,u^2]} \Bigg\{ \mathbb{E}\left\{C(\boldsymbol{\pi}_k, \boldsymbol{u})\right\} - g(\gamma)\bar{\tau}(\delta|k, \boldsymbol{\theta}_k) + \left[\sum_{m=1}^{M} \Pr(y_{k+1} = m|\boldsymbol{\theta}_k, u^2) V_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), \hat{o}_{k+1}(u^1), 1)\right] R\left(\delta|k\delta, \boldsymbol{\theta}_k\right) \tag{10}$$
$$+ V_\gamma(\boldsymbol{\pi}_0)(1 - R(k\delta, \boldsymbol{\theta}_k)) \Bigg\}, \qquad \forall \left\{\boldsymbol{\pi}_k = [k, \boldsymbol{\theta}_k, o_k, 1] \in \mathcal{S}\right\},$$

where $\mathbb{E}\left\{C(\boldsymbol{\pi}_k, \boldsymbol{u})\right\}$ and $\bar{\tau}(\delta|k, \boldsymbol{\theta}_k)$ are the expected cost and expected sojourn time at the current decision point given action $\boldsymbol{u} = [u^1, u^2]$. The 2nd line is the expected value function associated with the case that the system does not fail within the next period and the 3rd line is associated with the case where the system fails and it is brought back to state zero at the beginning of the next period. The optimal average cost satisfies the following optimality equation if the system is at the failure state at the kth decision epoch:

$$V_\gamma(\boldsymbol{\pi}_k) = \mathbb{E}\left\{C(\boldsymbol{\pi}_k, \gamma(\boldsymbol{\pi}_k))\right\} + V_\gamma(\boldsymbol{\pi}_0), \forall \left\{\boldsymbol{\pi}_k = [k, \boldsymbol{\theta}_k, o_k, \upsilon_k] \in \mathcal{S}, \upsilon_k = 0]\right\}, \tag{11}$$

where $\boldsymbol{\pi}_0 = (0, \boldsymbol{\theta}_0, \infty, 1)$. The expected cost at this state is the sum of the cost of replacement, the expected cost of failure, and the cost of warning (depending on weather or not a warning has been issued already). The optimal policy at this state is deterministic, that is replace immediately and issue warning if not been issued already. A pair $(g, V(\cdot))$, is said to be the optimal average cost solution if (10) holds true for all $\boldsymbol{\pi}_k \in \mathcal{S}$. Note that Equation (11) is the simplified version of (10) when the system is at the failure state. In the rest of this section, we develop the dynamic equations associated with (10) and simplify it based on possible actions $\boldsymbol{u} = [u^1, u^2]$. By definition, we know that there are only three actions available at the kth decision epoch when the warning process has not been issued yet, that is when the state of the system is $(k, \boldsymbol{\theta}_k, \infty, 1)$. These actions are (1) *do nothing* ($\boldsymbol{u} = [1, 1]$), (2) *warning immediately* ($\boldsymbol{u} = [0, 1]$), and (3) *replace and warning immediately* ($\boldsymbol{u} = [0, 0]$). The simplified version of the value functions of these cases are denoted by $V_\gamma^0(k, \boldsymbol{\theta}_k, \infty, 1)$, $V_\gamma^1(k, \boldsymbol{\theta}_k, \infty, 1)$, and $V_\gamma^2(k, \boldsymbol{\theta}_k, \infty, 1)$, respectively. Therefore, we get

$$V_\gamma(\boldsymbol{\pi}_k) = \min\left\{V_\gamma^0(\boldsymbol{\pi}_k), V_\gamma^1(\boldsymbol{\pi}_k), V_\gamma^2(\boldsymbol{\pi}_k)\right\}, \boldsymbol{\pi}_k = (k, \boldsymbol{\theta}_k, \infty, 1). \tag{12}$$

For the action do nothing with the value function $V_\gamma^0(\boldsymbol{\pi}_k)$, if the system fails within the next interval while it is in state i, then the total cost of replacement is $c_r + c_f^i$. Therefore the expected cost of replacement becomes $\sum_i c_r + c_f^i \theta_k^i \, \bar{R}(k\delta + \delta|k\delta, i)$ and it goes back to state $\boldsymbol{\pi}_0$ with the initial relative cost of $V_\gamma(\boldsymbol{\pi}_0)$. Therefore

$$[c_r + V_\gamma(\boldsymbol{\pi}_0)] \times \bar{R}(k\delta + \delta|k\delta, \boldsymbol{\theta}_k) + \sum_i c_f^i \theta_k^i \, \bar{R}(k\delta + \delta|k\delta, i)$$

is the expected replacement cost of the action *do nothing*, where $\bar{R}(k\delta + \delta|k\delta, \boldsymbol{\theta}_k)$ is the conditional probability of failure (1-reliability) within the next decision interval given $\boldsymbol{\theta}_k$. The expected cost of the warning process if failure occurs within the next inspection interval ($k\delta \le x \le k\delta + \delta$) given that the maintenance action is do nothing is $\int_{k\delta}^{k\delta+\delta} \mathscr{C}_l(x, x)\bar{R}(\mathrm{d}x|k\delta, \boldsymbol{\theta}_k)$. For the action warning immediately with the value function $V_\gamma^1(\boldsymbol{\pi}_k)$, the costs are the same except for the cost of warning, which becomes $\int_{k\delta}^{k\delta+\delta} \mathscr{C}_l(x, k\delta)\bar{R}(\mathrm{d}x|k\delta, \boldsymbol{\theta}_k)$. Finally for the action replace and warning immediately with the value function $V_\gamma^2(k, \boldsymbol{\theta}_k, \infty, 1)$, the expected cost of replacement process is only c_r

while the expected cost of the warning process becomes $\mathscr{C}_l(k\delta, k\delta) = c_s \times l$. We can now simplify all of the above to

$$V_\gamma^0(\boldsymbol{\pi}_k) = [c_r + V_\gamma(\boldsymbol{\pi}_0)] \times \bar{R}(k\delta+\delta|k\delta, \boldsymbol{\theta}_k) + \sum_i c_f^i \theta_k^i \, \bar{R}(k\delta+\delta|k\delta, i) + \int_{k\delta}^{k\delta+\delta} \mathscr{C}_l(x, x)\bar{R}(\mathrm{d}x|k\delta, \boldsymbol{\theta}_k) \tag{13}$$
$$- g(\gamma)\bar{\tau}(\delta|k, \boldsymbol{\theta}_k) + \left[\sum_{m=1}^{M} \Pr(y_{k+1} = m|\boldsymbol{\theta}_k, 1) \times V_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), \infty, 1)\right] R(k\delta+\delta|k\delta, \boldsymbol{\theta}_k),$$

$$V_\gamma^1(\boldsymbol{\pi}_k) = [c_r + V_\gamma(\boldsymbol{\pi}_0)] \times \bar{R}(k\delta+\delta|k\delta, \boldsymbol{\theta}_k) + \sum_i c_f^i \theta_k^i \, \bar{R}(k\delta+\delta|k\delta, i) + \int_{k\delta}^{k\delta+\delta} \mathscr{C}_l(x, k\delta)\bar{R}(\mathrm{d}x|k\delta, \boldsymbol{\theta}_k) \tag{14}$$
$$- g(\gamma)\bar{\tau}(\delta|k, \boldsymbol{\theta}_k) + \left[\sum_{m=1}^{M} \Pr(y_{k+1} = m|\boldsymbol{\theta}_k, 1) \times V_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), k, 1)\right] R(k\delta+\delta|k\delta, \boldsymbol{\theta}_k)\,,$$

$$V_\gamma^2(k, \boldsymbol{\theta}_k, \infty, 1) = c_r + \mathscr{C}_l(k\delta, k\delta) + V_\gamma(\boldsymbol{\pi}_0). \tag{15}$$

In the case that warning has been already issued at time d $(0 < d < k)$, there are only two possible maintenance actions: (1) do nothing and (2) replace immediately. Therefore we have

$$V_\gamma(\boldsymbol{\pi}_k) = \min\left\{V_\gamma^3(\boldsymbol{\pi}_k), V_\gamma^4(\boldsymbol{\pi}_k)\right\}, \boldsymbol{\pi}_k = (k, \boldsymbol{\theta}_k, d, 1), \tag{16}$$

where $V_\gamma^3(\boldsymbol{\pi}_k)$ and $V_\gamma^4(\boldsymbol{\pi}_k)$ respectively refer to the expected costs of *do nothing* ($\boldsymbol{u} = [1, 1]$) and *replace immediately* ($\boldsymbol{u} = [1, 0]$) assuming that the warning has already been issued at the dth decision epoch $(d < k)$. By simplifying (10), we get

$$V_\gamma^3(\boldsymbol{\pi}_k) = [c_r + V_\gamma(\boldsymbol{\pi}_0)] \times \bar{R}(k\delta+\delta|k\delta, \boldsymbol{\theta}_k) + \sum_i c_f^i \theta_k^i \, \bar{R}(k\delta+\delta|k\delta, i) + \int_{k\delta}^{k\delta+\delta} \mathscr{C}_l(x, d\delta)\bar{R}(\mathrm{d}x|k\delta, \boldsymbol{\theta}_k) \tag{17}$$
$$- g(\gamma)\bar{\tau}(\delta|k, \boldsymbol{\theta}_k) + \left[\sum_{m=1}^{M} \Pr(y_{k+1} = m|\boldsymbol{\theta}_k, 1) \times V_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), d, 1)\right] R(k\delta+\delta|k\delta, \boldsymbol{\theta}_k),$$

$$V_\gamma^4(\boldsymbol{\pi}_k) = c_r + \mathscr{C}_l(k\delta, d\delta) + V_\gamma(\boldsymbol{\pi}_0). \tag{18}$$

Note that the following holds true between the relative costs of replace immediately under the two cases of warning has already been generated and no warning has been issued as:

$$V_\gamma^2(k, \boldsymbol{\theta}_k, \infty, 1) = V_\gamma^4(k, \boldsymbol{\theta}_k, d, 1) + [\mathscr{C}_l(k\delta, k\delta) - \mathscr{C}_l(k\delta, d\delta)]\,. \tag{19}$$

Now, given the state of the system and the replacement policy (γ), we will show that it is possible to compare the cost of three possible options $V_\gamma^0(k, \boldsymbol{\theta}_k, \infty, 1)$, $V_\gamma^1(k, \boldsymbol{\theta}_k, \infty, 1)$, and $V_\gamma^2(k, \boldsymbol{\theta}_k, \infty, 1)$, and two possible actions $V_\gamma^3(k, \boldsymbol{\theta}_k, d, 1)$, and $V_\gamma^4(k, \boldsymbol{\theta}_k, d, 1)$ when $d > 0$ using some interesting properties of the value functions. After comparing all of these functions, we will end up with a control policy, which can determine what action is better depending on the state of the system.

5.2 Structural Properties of the Value Functions

This section presents how the set of controls can be determined for a policy γ with cost $g(\gamma)$. By mutually comparing the value functions associated with every two actions, we will show that at the kth decision point, the belief-dependent state $\boldsymbol{\pi}_k$ is sufficient to determine whether or not to (i) issue a warning and (ii) replace a degraded device. As the number of control actions is different depending on whether or not the warning has been issued, we perform our mutual comparison separately for the case of $0 < d < \infty$ and $d = \infty$. We introduced some important Lemmas on the structure of the value functions (see the Appendix), which will be used in Theorem 1-4 and the development of the optimal policy. The following four Theorems provide useful results for the relationships between the expected costs of possible maintenance actions for a policy γ when the state of the system is $\boldsymbol{\pi}_k$. The proofs for these Theorems are given in the Appendix.

Theorem 1 *The following holds true for any $\boldsymbol{\pi}_k = (k, \boldsymbol{\theta}_k, \infty, 1) \in \mathcal{S}$:*

$$V_\gamma^1(\boldsymbol{\pi}_k) \leq V_\gamma^2(\boldsymbol{\pi}_k), \quad \text{if } \Phi_\gamma^1(\boldsymbol{\pi}_k) \leq 0 \quad \text{and} \quad V_\gamma(\boldsymbol{\pi}_k) < V_\gamma^1(\boldsymbol{\pi}_k), \text{if } \Phi_\gamma^1(\boldsymbol{\pi}_k) > 0, \text{where}$$

$$\Phi_\gamma^1(\boldsymbol{\pi}_k) = \sum_i c_f^i \theta_k^i \, \bar{R}(k\delta + \delta | k\delta, i) - \mathscr{C}_l(k\delta + \delta, k\delta)\bar{R}(k\delta + \delta | k\delta, \boldsymbol{\theta}_k) - g(\gamma)\bar{\tau}(\delta | k, \boldsymbol{\theta}_k) \\ - c_s + \int_{k\delta}^{k\delta+\delta} \mathscr{C}_l(x, k\delta)\bar{R}(\mathrm{d}x | k\delta, \boldsymbol{\theta}_k). \tag{20}$$

This theorem shows one of the conditions under which warning and no replacement becomes better than warning and replacement. It also shows a condition under which warning and no replacement is not an optimal option. It can be observed from this Theorem that we can compare two actions under a policy γ, only if we know the average cost of this policy denoted by $g(\gamma)$. The rest of $\Phi_\gamma^1(\boldsymbol{\pi}_k)$ is fully known over time.

Theorem 2 *The following holds true for any $\boldsymbol{\pi}_k = (k, \boldsymbol{\theta}_k, \infty, 1) \in \mathcal{S}$:*

$$V_\gamma^0(\boldsymbol{\pi}_k) \leq V_\gamma^2(\boldsymbol{\pi}_k), \quad \text{if} \quad \Phi_\gamma^2(\boldsymbol{\pi}_k) \leq 0 \quad \text{and} \quad V_\gamma(\boldsymbol{\pi}_k) < V_\gamma^0(\boldsymbol{\pi}_k), \quad \text{if} \quad \Phi_\gamma^2(\boldsymbol{\pi}_k) > 0, \text{where}$$

$$\Phi_\gamma^2(\boldsymbol{\pi}_k) = \sum_i c_f^i \theta_k^i \, \bar{R}(k\delta + \delta | k\delta, i) - g(\gamma)\bar{\tau}(\delta | k, \boldsymbol{\theta}_k). \tag{21}$$

It is to be noted that this Theorem is compatible with the one-dimensional control policy developed in Ghasemi et al (2007) to compare the cost of do nothing with the cost of replacement, when there are only two possible actions and $c_f^1 = c_f^2 = \cdots = c_f^N$.

Theorem 3 *The following holds true any any $\boldsymbol{\pi}_k = (k, \boldsymbol{\theta}_k, \infty, 1) \in \mathcal{S}$:*

$$V_\gamma^0(\boldsymbol{\pi}_k) \leq V_\gamma^1(\boldsymbol{\pi}_k), \quad \text{if} \quad \Phi_\gamma^3(\boldsymbol{\pi}_k) \leq 0 \quad \text{and} \quad V_\gamma^0(\boldsymbol{\pi}_k) > V_\gamma^1(\boldsymbol{\pi}_k), \quad \text{if} \quad \Phi_\gamma^3(\boldsymbol{\pi}_k) > 0, \text{where}$$

$$\Phi_\gamma^3(\boldsymbol{\pi}_k) = \mathscr{C}_l(k\delta, k\delta)(1 - R(k\delta + \delta | k\delta, \boldsymbol{\theta}_k)) - \int_{k\delta}^{k\delta+\delta} \mathscr{C}_l(x, k\delta)\bar{R}(\mathrm{d}x | k\delta, \boldsymbol{\theta}_k) \\ + \left(\mathbb{E}^m\left[V_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), \infty, 1)\right] - \mathbb{E}^m\left[V_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), k, 1)\right]\right) R(k\delta + \delta | k\delta, \boldsymbol{\theta}_k). \tag{22}$$

This Theorem reveals the condition under which do nothing works better than warning immediately.

Theorem 4 *The following holds true for any $\boldsymbol{\pi}_k = (k, \boldsymbol{\theta}_k, d, 1) \in \mathcal{S}$:*

$$V_\gamma^3(\boldsymbol{\pi}_k) \leq V_\gamma^4(\boldsymbol{\pi}_k), \quad \text{if} \quad \Phi_\gamma^4(\boldsymbol{\pi}_k) \leq 0 \quad \text{and} \quad V_\gamma^3(\boldsymbol{\pi}_k) > V_\gamma^4(\boldsymbol{\pi}_k), \quad \text{if} \quad \Phi_\gamma^4(\boldsymbol{\pi}_k) > 0, \text{where}$$

$$\Phi^4_\gamma(\boldsymbol{\pi}_k) = \sum_i c^i_f \theta^i_k \bar{R}(k\delta+\delta|k\delta,i) - g(\gamma)\bar{\tau}(\delta|k,\boldsymbol{\theta}_k) + [\mathscr{C}_l(k\delta+\delta,d\delta)R(k\delta+\delta|k\delta,\boldsymbol{\theta}_k) - \mathscr{C}_l(k\delta,d\delta)] + \int_{k\delta}^{k\delta+\delta} \mathscr{C}_l(x,d\delta)\bar{R}(\mathrm{d}x|k\delta,\boldsymbol{\theta}_k). \quad (23)$$

This Theorem reveals the condition under which do nothing is better than replace immediately, when the warning has already been issued.

Remark 2 For all $\boldsymbol{\pi}_k = (k,\boldsymbol{\theta}_k,\infty,1) \in \mathcal{S}$, $\Phi^1_\gamma(\boldsymbol{\pi}_k) \leq \Phi^2_\gamma(\boldsymbol{\pi}_k)$.

Proof The proof is given in the Appendix. This property will be used later as part of the development of the optimal policy. For example, if $\Phi^1_\gamma(\boldsymbol{\pi}_k) > 0$, then according to this remark $\Phi^2_\gamma(\boldsymbol{\pi}_k) > 0$, and therefore from Theorems 1-2, we have $V_\gamma(\boldsymbol{\pi}_k) = V^2_\gamma(\boldsymbol{\pi}_k)$, which means warning and replace immediately is the best action.

5.3 Form of the Control Policy

Based on the mutual comparison we made between the relative costs of each two actions in the previous section, we can develop a control policy with eight control rules as given in Remark 3. These control rules are the output of comparing the functions at each decision point $k\delta$, $k \in \mathbb{N}$ using Theorems 1-4. These control rules determine which option to choose based on the state of the system.

Remark 3 The optimal joint warning-replacement policy γ is characterized by 4 critical control measures $\Phi^1_\gamma(\boldsymbol{\pi}_k), \Phi^2_\gamma(\boldsymbol{\pi}_k), \Phi^3_\gamma(\boldsymbol{\pi}_k), \Phi^4_\gamma(\boldsymbol{\pi}_k)$ for $\boldsymbol{\pi}_k = (k,\boldsymbol{\theta}_k,o_k,\upsilon_k)$ as defined in Eqs. (20)-(23). In particular, at the kth decision epoch with state-belief of $\boldsymbol{\pi}_k \in \mathcal{S}$, the following 8 control rules (CR), referred to as CR1-CR8 can be used for warning generation and replacement decision-making:

$$\gamma(\boldsymbol{\pi}_k) = \begin{cases} (0,0), & \text{if } \upsilon_k = 0, o_k = \infty \rightarrow \text{CR1} \\ (1,0), & \text{if } \upsilon_k = 0, o_k \neq \infty \rightarrow \text{CR2}, \\ (0,0), & \text{if } \upsilon_k = 1, o_k = \infty, \Phi^1_\gamma(\boldsymbol{\pi}_k) > 0 \text{ , } \Phi^2_\gamma(\boldsymbol{\pi}_k) > 0 \rightarrow \text{CR3}, \\ (0,1), & \text{if } \upsilon_k = 1, o_k = \infty, \Phi^1_\gamma(\boldsymbol{\pi}_k) < 0 \text{ , } \Phi^2_\gamma(\boldsymbol{\pi}_k) > 0 \rightarrow \text{CR4} \\ (1,1), & \text{if} \upsilon_k = 1, o_k = \infty, \ \Phi^2_\gamma(\boldsymbol{\pi}_k) < 0 \text{ , } \Phi^3_\gamma(\boldsymbol{\pi}_k) \leq 0 \rightarrow \text{CR5} \\ (0,1), & \text{if } \upsilon_k = 1, o_k = \infty, \ \Phi^2_\gamma(\boldsymbol{\pi}_k) < 0 \text{ , } \Phi^3_\gamma(\boldsymbol{\pi}_k) > 0 \rightarrow \text{CR6} \\ (1,1), & \text{if } \upsilon_k = 1, o_k = d \neq \infty, \Phi^4_\gamma(\boldsymbol{\pi}_k) \leq 0 \rightarrow \text{CR7} \\ (1,0), & \text{if } \upsilon_k = 1, o_k = d \neq \infty, \Phi^4_\gamma(\boldsymbol{\pi}_k) > 0 \rightarrow \text{CR8}, \end{cases}. \quad (24)$$

Proof The proof can be directly obtained from Theorems 1-4 and Remark 2. Note that, the device is replaced at failure regardless of its state (CR1-CR2). In such a condition, if the warning has not been issued yet, it is issued immediately (CR1). When the system is available and the warning has not been issued yet, there are 4 control rules (CR3-CR6), for which the control actions are directly determined from Theorems 1-2. On the other hands, when the warning has already been issued at the dth decision point, there are two control rules (CR7-CR8), which are originated from Theorems 3-4. It can be seen by carefully looking at Eq. (24) that the eight control rules correspond to eight mutually exclusive conditions, thus, the system is at exactly one of these conditions at any point of time.

The above control rules can be visualized as simple control charts, which are more intuitive for decision makers. From Remark 3, $T_{1,\gamma}$ and $T_{2,\gamma}$ can be found as

$$T_{1,\gamma} = \min\left\{\zeta, \inf\left\{k\delta \in \mathbb{R}^+ : \Phi^2_\gamma(k,\boldsymbol{\theta}_k,\infty,1) > 0 \vee \Phi^2_\gamma(k,\boldsymbol{\theta}_k,\infty,1)\Phi^3_\gamma(k,\boldsymbol{\theta}_k,\infty,1) < 0\right\}\right\}, \quad (25)$$

$$T_{2,\gamma} = \min\left\{\zeta, \inf\left\{k\delta \in \mathbb{R}^+ : \Phi_\gamma^2(k, \boldsymbol{\theta}_k, \infty, 1) > 0 \vee \Phi_\gamma^4(k, \boldsymbol{\theta}_k, d, 1) > 0\right\}\right\}. \tag{26}$$

The above equations verify that the warning will be issued as soon as one of the conditions given in Eq. (25) is satisfied or the system fails, whichever occurs first. Also, the system will be replaced as soon as one of the two conditions given in Eq. (26) is satisfied or the system fails, whichever occurs first.

Remark 4 The introduced control policy has a monotonic behavior for $k_1 \le k_2$ as follows:

$$\gamma_2(k_1, \boldsymbol{\pi}_{k_1}, o_{k_1}, \upsilon_{k_1}) \le \gamma_2(k_2, \boldsymbol{\pi}_{k_2}, o_{k_2}, \upsilon_{k_2}).$$

Proof The above Remark states that the actions chosen at future decision epochs are not weaker than the action chosen at the current decision epoch. For example, if the action at the current time is to *replace immediately* but we let the system continue working, then the action suggested at any point in future will also be to replace immediately regardless of the state of the system given that the system does not self heal. This property can be proved by comparing the control indices provided in Remark 3. It can be easily shown that Φ_γ^1, Φ_γ^2, and Φ_γ^4 are monotonically nondecreasing in k and $\boldsymbol{\theta}$, for all $\boldsymbol{\theta} \in \Theta$ under the monotonicity assumption for the conditional reliability function. We omit the proof for its simplicity.

5.4 Finding the Optimal Policy

It is clear that the only unknown element of $\Phi_\gamma^1 - \Phi_\gamma^4$ is $g(\gamma)$, which is the average cost of policy γ. Therefore, in order to find the optimal policy, we should find the cost of the optimal policy first. In order to find the optimal strategy, one may transform it to an equivalent parameterized stochastic control problem with an additive objective function as described in Aven and Bergman (1986) and Kim and Makis (2013) or use other standard approaches, such as policy iteration and value iteration (Bertsekas and Tsitsiklis, 1996). The former approach is known as the lambda-minimization technique and its theory was developed in the excellent paper of Aven and Bergman (1986). Based on this technique, denoting M^T and S^T as the expected cost and cycle of a policy, λ^* which is the solution of $C_\lambda^T = M^T - \lambda S^T$ is the optimal expected average cost for the stochastic control problem. In this paper, we employed the policy iteration approach as it gives nice structural properties and easier to implement for our framework. Policy iteration, which is a well-known algorithm used to solve dynamic programming problems, is an iterative procedure that operates in a way that an initial policy is improved until no further improvement is possible. Policy iteration has two main steps, referred to as policy evaluation and policy improvement. Starting from an initial policy, we compute the value functions by solving the dynamic system of equations (policy evaluation step). Then, we find the new set of actions and a new policy using the updated value functions (policy improvement step). The policy iteration stops when further improvement is impossible. Directly applying the policy iteration for solving the partially observable semi-Markov process is computationally intractable due to the number of states and the size of the system of dynamic equations. To be able to apply the policy iteration technique, we can modify its steps in the sense that the policy improvement and policy evaluation steps can be performed without having to deal with solving directly the associated dynamic equations. We have already shown all the steps needed for the policy improvement in §5.2. In §5.4.1, we show how the average cost function can be calculated for a given policy (policy evaluation). We summarize our results in §5.6.

5.4.1 Steps for Policy Evaluation

Now that the structure of the warning-replacement policy is determined (policy improvement step), the remaining step of the policy iteration is to perform the policy evaluation step, which evaluates a policy and finds a cost associated with it accordingly. In order to do this, we define the terms $\bar{W}_\gamma(\boldsymbol{\pi}_k)$, $\bar{Q}_\gamma(\boldsymbol{\pi}_k, i)$, and $\bar{C}_\gamma(\boldsymbol{\pi}_k)$, which respectively denote the expected remaining time to replacement, the expected probability of failure replacement at state i, and the expected cost of the warning process, given the current state of the system $(k, \boldsymbol{\theta}_k, o_k, \upsilon_k)$ and policy γ. The average cost function given in (9) can then be evaluated given that,

$$\mathbb{E}(\min(\zeta, T_{2,\gamma})) = \bar{W}_\gamma(\boldsymbol{\pi}_0), \mathbb{E}\Big(\Pr(\zeta \leq T_{2,\gamma}, Z(L) = i)\Big) = \bar{Q}_\gamma(\boldsymbol{\pi}_0, i), \text{and}$$

$$\mathbb{E}\left(\mathscr{C}_l(\min(\zeta, T_{2,\gamma}), \min(\zeta, T_{1,\gamma}))\right) = \bar{C}_\gamma(\boldsymbol{\pi}_0).$$

Let use also define $D_\gamma(\boldsymbol{\pi}_k, m, u_k^1)$ as the next state of the system given that the next observation is m, the device will survive until time $(k+1)\delta$, and the warning control is u_k^1. Then we have

$$D_\gamma(\boldsymbol{\pi}_k, m, u_k^1) = (k+1, \hat{\theta}^j_{k+1}(m), \hat{o}_{k+1}(u_k^1), 1),$$

and therefore the following holds true:

$$\bar{W}_\gamma(\boldsymbol{\pi}_k) = \gamma_2(\boldsymbol{\pi}_k)\left(\bar{\tau}(\delta|k, \boldsymbol{\theta}_k) + R(k\delta+\delta|k\delta, \boldsymbol{\theta}_k) \sum_{m=1}^{M} \Pr(y_{k+1} = m|\boldsymbol{\theta}_k, 1)\ \bar{W}_\gamma\big(D_\gamma(\boldsymbol{\pi}_k, m, u_k^1)\big)\,\gamma_2\big(D_\gamma(\boldsymbol{\pi}_k, m, u_k^1)\big)\right), \tag{27}$$

$$\bar{Q}_\gamma(\boldsymbol{\pi}_k, i) = \gamma_2(\boldsymbol{\pi}_k)\left(\theta_k^i \bar{R}(k\delta+\delta|k\delta, i) + R(k\delta+\delta|k\delta, \boldsymbol{\theta}_k) \sum_{m=1}^{M} \Pr(y_{k+1} = m|\boldsymbol{\theta}_k, 1)\ \bar{Q}_\gamma\big(D_\gamma(\boldsymbol{\pi}_k)\big)\,\gamma_2\big(D_\gamma(\boldsymbol{\pi}_k, m, u_k^1)\big)\right). \tag{28}$$

The above equations imply that when the replacement control at the kth decision epoch is to *replace immediately* (i.e.,$\gamma_2(\boldsymbol{\pi}_k) = 0$), then the expected time to failure and expected probability of failure within a cycle are zero. However, if the decision is *do nothing*, then these elements depend on the posterior evolution of states, observation signals, and the actions associated with them. For the expected cost of the warning process, we should note that if the decision at the kth decision epoch is to *replace immediately*, then the cost depends on whether or not the warning has already been issued (i.e., $\mathscr{C}_l(k\delta, \hat{o}_{k+1}(u_k^1))$). However, if the decision is *do nothing*, the expected cost of warning equals the expected cost of warning given that the system fails during the next interval, plus the expected cost of warning if the system survives until time point $k\delta + \delta$. Thus

$$\begin{aligned}\bar{C}_\gamma(\boldsymbol{\pi}_k) &= (1-\gamma_2(\boldsymbol{\pi}_k))\mathscr{C}_l(k\delta, \hat{o}_{k+1}(u_k^1)) + \gamma_2(\boldsymbol{\pi}_k) \\ &\times \left(\int_{k\delta}^{k\delta+\delta} \mathscr{C}_l(x, \hat{o}_{k+1}(u_k^1))\bar{R}(\mathrm{d}x|k\delta, \boldsymbol{\theta}_k) + R(k\delta+\delta|k\delta, \boldsymbol{\theta}_k) \sum_{m=1}^{M} \Pr(y_{k+1} = m|\boldsymbol{\theta}_k, 1)\ \bar{C}_\gamma\Big(D_\gamma(\boldsymbol{\pi}_k, m, u_k^1)\Big)\right).\end{aligned} \tag{29}$$

In order to find the above measures for any policy, a backward recursive procedure should be designed in the sense that it is computationally feasible. To do this, we can first find the time point at which the reliability function is approximately zero as the start of the backward process. This is theoretically the time point that the warning and replacement immediately are imposed regardless of the state of the system. This time point can be considered as a starting point of the backward recursive procedure. Algorithm 1 given in the Appendix describes the details of the Backward Recursive Process to find $\bar{Q}$, $\bar{W}$, and $\bar{C}$ for a given policy γ.

5.5 Shrinking the State Space from $\mathcal{S}$ to $\mathcal{S}'$

As the infinite horizon solution of the described POSMDP with continuous state space is computationally intractable, we discretize the belief space Θ, so that the state space becomes finite and discrete based on the discretization step σ $(0 \leq \sigma \leq 1)$ as shown below

$$\tilde{\Theta}_{\sigma} = \left\{ \left\{ a^1, \cdots, a^N \right\}; 0 \leq a^i \leq 1, a^i = q^i \sigma, \forall q^i \in \mathbb{N}_0, \forall\, i \in \mathscr{E}, N, \sum_{i=1}^{N} a^i = 1 \right\}.$$

It is important to note that some of the above states might be totally unreachable. The belief state $\boldsymbol{\theta}_k = \{a^1, a^2, \ldots, a^N\}$ is unreachable if there exists no observation sequence $\{y_1, \cdots, y_k\}$ that yields $\boldsymbol{\theta}_k$. One can run a simple empirical forward process to simulate multiple run-to-failure CM sequences in order to compute the reachable belief space before running the policy iteration algorithm. Another possible way to decrease the size of the state space is to define T_{max} as the time point at which the system is definitely replaced. Therefore, all states $\boldsymbol{\pi}_k$, where $k\delta > T_{max}$ are not reachable when the degrading system is under control.

5.6 Summary of the Development of the Optimal Control Policy

The following summarizes the steps needed to develop the optimal policy. The decision process has two phases, which are training phase and implementation phase. In the training phase, the structure of the optimal warning-replacement policy and its associated cost are determined (Policy Improvement and Policy Iteration). In the implementation phase, using real-time condition monitoring data, the trained decision policy is used for online maintenance decision-making using CM data. The summary of steps for each phase is illustrated in Algorithms 2-3. These algorithms generate an improving sequence of policies and terminates with an optimal policy.

Algorithm 2: Policy Iteration Algorithm to Develop a Warning-Replacement Policy

Step 1. *Initialization.* Set $n := 0$ and choose a tolerance limit $\epsilon > 0$. Define $\gamma^{(0)}$ as the policy do-nothing with its associated cost $g(\gamma^{(0)})$ given in the upper bound in Remark 1.

Step 2. *Policy Improvement.* For each $\boldsymbol{\pi} \in \mathcal{S}'$ (where $\mathcal{S}'$ is the discretized state space), determine the control actions using Remark 3. The resulting optimal actions for each $\boldsymbol{\pi} \in \mathcal{S}'$ obtained from Remark 3, yields the new policy $\gamma^{(n+1)}$.

Step 3. *Policy Evaluation.* Find the cost associated with policy $\gamma^{(n+1)}$ using the results in §5.4.1.

Step 4. *Iteration*: If $|g(\gamma^{(n+1)}) - g(\gamma^{(n)})| < \epsilon$, then $\gamma^{(n+1)}$ is the optimal replacement policy with associated cost $g^* = g(\gamma^{(n+1)}), \gamma^* = \gamma^{(n+1)}$. Otherwise, set $n := n + 1$ and move back to Step 2.

Algorithm 3: Implementation Steps for Online Decision-Making

Step 1: Set $k := 0$. The decision action at this point (time zero with state $\boldsymbol{\pi}_0$) is *do nothing* (DN).

Step 2: Set $k := k + 1$. Collect the condition monitoring signal (y_k) and update $(\boldsymbol{\pi}_k)$.

Step 3: Determine the set of control actions using Remark 3. If $\gamma_2(\boldsymbol{\pi}_k) = 0$, then replace the system and move back to Step 1, otherwise move back to Step 2. The summary of the decision process is shown in Figure 3.

6 Search for Optimal Static Decision Variables

As stated before, our static decision variables can be determined offline. In the rest of this section, the steps for finding the solution of the two static decision variables, namely, monitoring interval and lead time are discussed. Let us assume that the unit cost of condition monitoring is $c_{m1}(\delta)$ per unit of operation time and the cost per monitoring signal is $c_{m2}(\delta)$, where δ is the time interval between two monitoring points. It should be pointed out that for many systems, $c_{m1}(\delta)$ and $c_{m1}(\delta)$ may be negligible compared to the costs of replacement and failure.

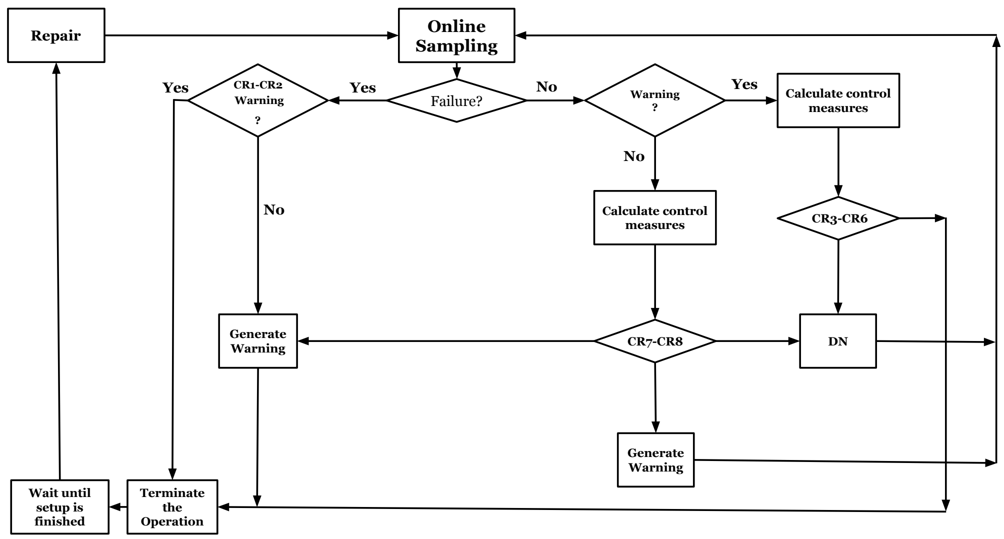

Fig. 3: Decision-making flowchart for warning and replacement (Implementation Phase)

Note that the costs of ordering and monitoring were not included in the average cost function as they do not influence the dynamic variables $T_{1,\gamma}$ and $T_{2,\gamma}$. The full average cost function per time unit can be rewritten as follows:

$$g(\gamma, l, \delta) = g(\gamma) + c_{m1}(\delta) + \frac{c_{m2}(\delta)}{\delta} + \frac{c_o(l)}{\mathbb{E}(\min(\zeta, T_{2,\gamma}))},$$

where $c_{m1}(\delta) + \frac{c_{m2}(\delta)}{\delta}$ is the cost of condition monitoring per unit of time, $\frac{c_o(l)}{\mathbb{E}(\min(\zeta, T_{2,\gamma}))}$ is the cost of ordering per unit of time, and $g(\gamma)$ is the regular average cost of control given in (9).

6.1 Algorithm to Find the Optimal Monitoring Interval

In this section, we propose a simple approach to compare policies with different observation intervals when $c_{m1}(\delta)+c_{m2}(\delta) > 0$. Consider situations where instead of monitoring every δ units of time, the signals are collected every Δ units of time where $\Delta = j\delta, j \in \mathbb{N}$. While the decision can still be made at $h\delta, h \in \mathbb{N}$, the observation vector is updated only every Δ units of time. As the number of transitions within each observation interval Δ can be more than one, the distributions of the hazard rate, reliability, and sojourn time are different from the case with smaller monitoring intervals δ. Let us define y' as a time-dependent dummy observation output with the following stochastic relationship with the Z process:

$$\Pr(Y_k = y'|Z_k = i) = b_i(y') = 1, \qquad \forall\{k : k\delta \neq h\Delta, h \in \mathbb{N}\}.$$

The above relationship states that at any decision point that is not a signal collection point, the observation process generates an artificial output y', which has the same distribution among all states. Therefore, this observation does not add any information to decision makers and is stochastically equivalent to having no observation signal. A sample sequence of observations when the observation interval equals Δ is $\{y', y', \cdots, y_\Delta, y', y', \cdots\}$. Now, we can simply assume that the system is still under monitoring at decision epochs $\delta, 2\delta, ...$, but the monitoring process at $h\delta \neq \Delta, h \in \mathbb{N}$, has no cost and it generates y', which does not add any useful information for decision-making. In order to incorporate this in the development of the optimal policy, we need to modify the observation matrix, in a sense that it becomes a time-varying observation matrix $\mathbf{B}(t)$ as follows:

$$\mathbf{B}(t) = \begin{cases} [\mathbf{B_0} \quad \mathbf{0}_{N\times 1}] & \forall\{t : t = h\varDelta, h \in \mathbb{N}\} \\ \\ [\mathbf{0}_{N\times M} \quad \mathbf{1}_{N\times 1}] & \forall\{t : t = h\delta, t \neq h\varDelta, h \in \mathbb{N}\} \end{cases}, \tag{30}$$

where $\mathbf{B}_0$ is the original observation matrix. According to (30), at any observation point $t = h\varDelta, h \in \mathbb{N}$, the observation probability matrix works the same as the original observation matrix, however, at other monitoring points, the system generates output y' with probability 1. In the development of the optimal policy, we need to replace $b_i(j)$ with its associated value from the above equation. The following summarizes how to find the optimal inspection interval $\varDelta$ from a set of alternatives in a set $\boldsymbol{\Delta} = \{\varDelta_1, \dots, \varDelta_F\}$:

Step 1. Update $\mathbf{B}(t)$ using (30).

Step 2. Find the optimal warning-replacement cost $g(\gamma^*, l, \varDelta)$.

Now by comparing the cost of each option, the optimal interval, which is the one with the minimum cost, can be found as $\varDelta^* = \arg \min_{\varDelta \in \boldsymbol{\Delta}} g(\gamma^*, l, \varDelta)$. Given the steps described above, our model can be used (i) to find the optimal inspection/monitoring interval from a set of alternatives, (ii) to compare a system under condition monitoring with a system under inspection and with a system under no inspection and monitoring, and (iii) to analyze cases where monitoring/inspection intervals are large (not necessarily equal to the decision interval). Similar to any other type of investment, having an online monitoring system needs to be evaluated economically in order to find out whether or not the investment is worth it and how suitable it is in a given application. The factors that need to be taken into account are the cost saving per unit time using the online monitoring tool, the investment cost of this system, the maintenance cost associated with this system, and other important non-monetary factors. In order to figure out whether or not a condition monitoring framework is worth it, we can compare the full average cost associated with this system with the one obtained from a fully observable degradation process with inspection. This will be shown by a numerical example in §7.

6.2 Optimal Lead Time

The assumption we have had so far was the lead time (whether it related to maintenance setup or spare part ordering) is fixed and already known. In this subsection, we explain how we can find the optimal lead time only if there is a flexibility to reduce it if we pay more. A useful application of this model is when emergency ordering or emergency maintenance setup, which has a shorter lead time and a higher cost than a regular one is possible. The search for the optimal lead time is very straightforward. If the cost of ordering is zero, then the ideal lead time is its minimum possible value. However, if the cost of ordering for lower lead times is greater than its cost for higher lead times, then the search for the optimal lead time is reasonable. Let $\mathbf{L} = \{l_1, l_2, \dots, l_I\}$ be set of alternatives for the lead time with associated cost $\{c_o(l_1), c_o(l_2), \dots, c_o(l_I)\}$. Then the optimal lead time is $l^* = \underset{l \in \mathbf{L}}{argmin}\, g(\gamma^*, l, \delta)$. It should be pointed out that if the cost of early warning is zero, the optimal warning policy is to issue warning immediately regardless of l. Similarly, if the cost of late warning is zero (which is often not the case in practice), then the optimal warning policy is to issue warning at failure regardless of l. In addition to the above, decision makers may choose to run an emergency order or emergency maintenance setup when $T_{1,\gamma} = T_{2,\gamma}$. This is possible only if there exists a maintenance setup option with lead time $x < l$ and cost $c_o(x)$ under which the the cost of warning and ordering is less than the original cost of warning and ordering based on lead time l, that is $c_s \times x + c_o(x) < c_s l + c_o(l)$. If an $x < l$ satisfying the above inequality exists, decision makers may run an emergency setup/order.

6.3 Special Cases

This section provides some useful results for important special cases with respect to the observation process.

I. Completely observable process: Since a degrading system may be completely observable over time (e.g., inspections are periodically held every Δ time units, during which the true state of the system is identified), we briefly explain how our model can be used for such cases. The first step is to assume that there exists an observation process with outputs $\{1,\dots, N\}$, which perfectly reveal the process Z. In other words, the observation probability matrix is an identity matrix with size N, i.e., $\mathbf{B}(t) = \mathbf{I}_N$. Now, we use this matrix as an input to develop the optimal policy. In the implementation phase, we generate artificially an observation process based on $\mathbf{B}$ and the observed true states and use it for decision-making. This result can also be used to compare a system under condition monitoring with a system under full inspection. Remember, the system can still be controlled every δ time units. A sample sequence of observations follows the form $\{y', y', \cdots, Z(\Delta), y', y', \cdots\}$.

II. Missing points: Dealing with missing point in our model is very simple. We can assume a dummy observation value y' with the corresponding column vector $\mathbf{1}_{N\times 1}$ in $\mathbf{B}(t)$ similar to (30). Then, we can simply replace all missing points with an equivalent dummy signal y' that does not add any information for decision-making.

III. Large monitoring intervals: One of the common assumptions made in the literature for monitoring/inspection interval is that it is small enough so that at most one transition may occur within each interval. Our model can still be used for decision-making even when the monitoring interval is large (Δ). This can be done by adding a dummy observation and updated $\mathbf{B}$, we can implement the decision process by assuming that the decision interval is still δ, but the observation interval is $\Delta\,(\delta \leq \Delta)$. A sample sequence of observations follows the form $\{y', y', \cdots, y_\Delta, y', y', \cdots\}$.

IV. Outliers: As the observation process is evolved according to sensor measurements, it is always possible that we observe an output deviated markedly from other observations in the sample. We detect outliers by comparing the posterior observation distribution given in (7) with a pre-defined critical point η, $0 < \eta < 1$. In other words, if the probability of observing the outcome y_k is less than the critical point η, then the observed outcome y_k is considered to be outlier and can be treated as a missing point (see II in Special Cases).

6.4 Shortcoming of the Proposed Model

The results provided in this paper are subject to three main shortcomings: (i) scalability, (ii) multiple CM signals, and (iii) model structure. In terms of the scalability and computational complexity, our model may not perform well if the number of states is very large. We should point out that since the policy iteration steps are conducted offline and most multi-state frameworks have few states that represent the overall health of the system, this issue is not of a significant drawback of our work. To deal with such an issue, we have two options: (i) use the results given in Section 5.5 to shrink the state space, and (ii) run the model training phase for a longer time. If we have a very large number of states (which is often not the case in the reliability domain) and none of the two given solutions works, then our approach may not be scalable. The second drawback is the ability of our model to deal with system where multi-dimensional condition monitoring signals exist and feature fusion models are not effective to combine them to one signal. Although this is not addressed in this paper, one can extend the results to cover multi-dimensional observation process without much effort as long as the assumptions made remain valid. The other shortcoming is the assumption that the structure of the degradation process and observation processed are already trained with data and the trained structure is a reasonable representative of the dynamics of the system. If one or more of the assumptions of the model are violated, then our model may give misleading results.

7 Numerical Experiments

In this section, a set of numerical experiments is given to show how the proposed control model can be used to employ condition monitoring data for decision-making of two types: (1) when to issue a warning and (2) when to stop the operation for replacement. We also illustrate the possible benefit of our model with respect to traditional models. Consider a three-state deterioration system with the following transition probability matrix ($\mathbf{P}$) and state observation matrix ($\mathbf{B}$):

$$\mathbf{P} = \begin{bmatrix} 0.80 & 0.19 & 0.01 \\ 0 & 0.77 & 0.23 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0.60 & 0.30 & 0.05 & 0.05 & 0 \\ 0.10 & 0.20 & 0.40 & 0.20 & 0.10 \\ 0 & 0.05 & 0.05 & 0.30 & 0.60 \end{bmatrix}.$$

The failure rate is $\lambda(t, Z(t)) = (4.5/18)(t/18)^{3.5} \exp(1.6\,(Z(t) - 1))$, where $Z(t) \in [1, 2, 3]$. Except for the failure state, the degradation level of the system is assumed to be not observable and only stochastically related with the actual level of degradation defined through $\mathbf{B}$. For simplicity, we have assumed that the failure cost is constant over all states (i.e., $c_f^1 = c_f^2 = \cdots = c_f^N = c_f$). We should point out here that in this paper, we have assumed that all parameters associated with the degradation and observation processes are fully known. Thus, parameter estimation is out of the scope of this work. Interested readers are referred to Ghasemi et al (2010) and references therein for parameter estimation. For all our computational tests, we used Matlab on an Intel Corel i7, 2.7 GHz with 16GB RAM in all of numerical experiments. We discretized the belief state with $\sigma = 0.01$. We set $T_{max} = 30$, as the point at which the probability of failure is almost 1.

7.1 Development of the Optimal Policy

We first set $c_r = 10$, $c_h = 0.1$, and $l = 2$, and found the optimal cost function for several combinations of c_f/c_r, and c_s/c_h, as given in Table 1. From this table, we can observe that as expected, the optimal average cost function increases when c_f/c_r and c_s/c_h increase, but the rate of change is much more sensitive to c_f/c_r. The optimization procedure starts with an initial policy of *do nothing* with the average cost calculated from Remark 1 (upper bound of g) and ends when the policy iteration cannot further improve it. In almost all cases, the optimal solution was found within the first 6 iterations. Figure 4 shows the output of the policy iteration for a case where $c_f/c_r = 5$, and $c_s = c_r$. For this example, the average cost function is 1.93, the expected probability of a failure replacement is 3.87%, and the expected replacement cycle is 6.2. The lower bound shown here is $\frac{c_r}{\mathbb{E}(\zeta)-\delta}$, which is the average cost per unit time if the policy yields no failure replacement, no late or early warnings, and yields maximum possible operating time.

Table 1: Optimal average cost for various combinations of c_f/c_r and c_s/c_h

c_s/c_h	c_f/c_r			
	0	1	5	10
1	1.04	1.50	1.93	2.17
10	1.05	1.51	1.95	2.18
50	1.06	1.53	1.95	2.19
100	1.09	1.89	1.96	2.19

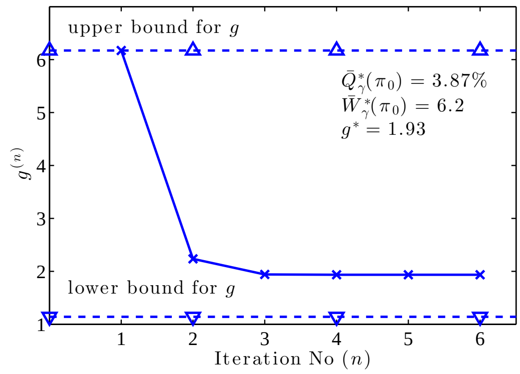

Fig. 4: An example of the development of the optimal policy using policy iteration

7.2 Construction of the Optimal Control Charts

As discussed earlier, the proposed control policy can be used for the joint optimization of the warning time and the replacement time using online CM measurements. To do this, one can convert the control indices given in Remark 3 to simple control charts to be used for online decision-making. We do this for one sample with a lifetime of 9.37. To generate the decision process, we simulate the degradation process and its associated observation process and then based on that calculate the state of the system at each decision epoch. As discussed earlier, when the warning has not been issued, we have three possible options of do nothing, warning immediately, and warning and replacement. Therefore, $\Phi^1_{\gamma^*}(\boldsymbol{\pi}_k)$ - $\Phi^3_{\gamma^*}(\boldsymbol{\pi}_k)$ can be used for decision making as shown in Figure 5. However, after the warning is issued, the operator has only two options, whether to do nothing or replace immediately. Here, the control index $\Phi^4_{\gamma^*}(\boldsymbol{\pi}_k)$ should be used for this purpose as given in Figure 6. From these control charts, we observe that for this sample, the optimal warning time and the optimal replacement time are respectively 6 and 9. Therefore, we have a preventive replacement, and one unit early warning time (the time between warning and replacement is 3, but the lead time is 2). For this example, the operators could potentially prevent failures and issue warning reasonably on time (one unit early). These control charts are very practical as they are intuitively very easy to understand and monitor and can be used directly for decision-making.

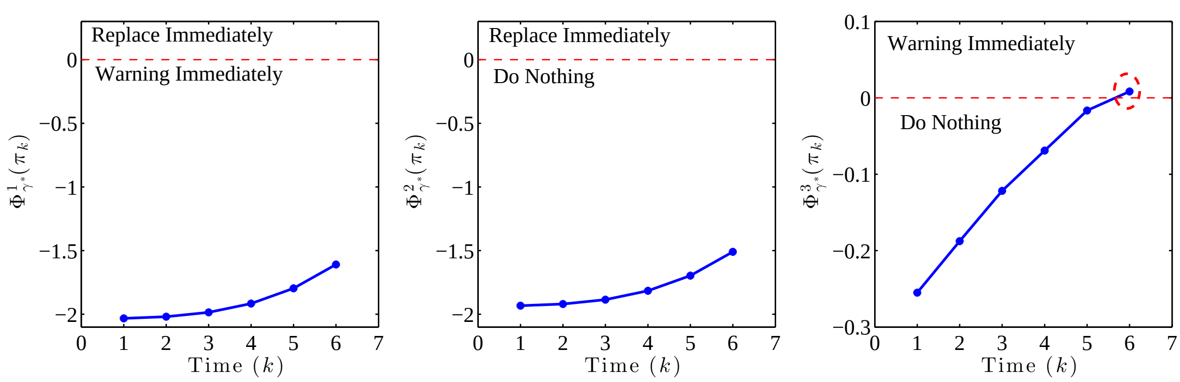

Fig. 5: Control charts before the warning is being issued

7.3 Comparison with Other Policies

In this subsection, we compare through several numerical examples and parameter settings the performance of our proposed joint optimal policy with other policies used in the literature for the same purpose. We should point out that our goal here is not to prove that our model outperforms others (this has already been proved as our model provides the optimal policy), but to show that how a joint optimization of decision variables using online CM data

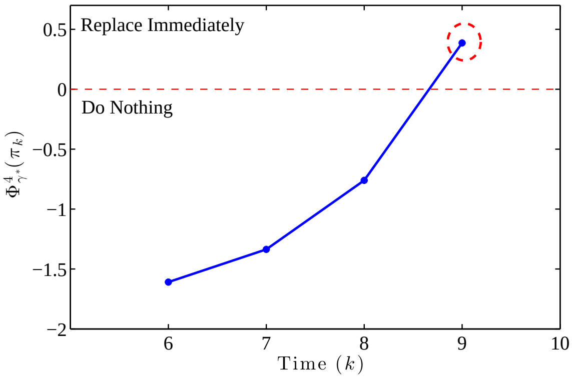

Fig. 6: The control chart after the warning is being issued

can help in preventing unexpected failures and unnecessary maintenance actions. Since our paper investigated the optimal policy, it should give better or equal results compared to any other policies. We compare our model with three other policies, Do-nothing policy, (T_1, T_2) policy, and sequential policy. The Do-nothing policy forces the warning and replacement time to be equal to the failure time. Therefore, the system is always subject to a failure replacement and too late warning. The (T_1, T_2) policy is the traditional age-based policy, which does not take into account the condition monitoring data and works based on the distribution of the time to failure (Armstrong and Atkins, 1996). Based on this policy, the system is replaced every T_2 time units or at failure whichever occurs first, and the warning is generated at time T_1 or the replacement time, whichever occurs first ($T_1 \leq T_2$). The value of T_1 and T_2 can be obtained analytically or through simulation. A dynamic sequential policy is a policy, which (i) uses CM data to update the cost function over time, (ii) first optimizes the replacement time, and then (iii) finds the warning time accordingly. One recent example of such a policy is the one of Elwany and Gebraeel (2008), in which the replacement time is first found by optimizing the cost function using the revised remaining useful life distribution updated continuously according to real-time condition-based degradation signals. In other words, at each decision epoch, without considering the ordering process and its associated cost, the replacement time is determined. Then the computed optimal replacement time is incorporated into a full cost function to compute the optimal inventory ordering time. The sequential policy we consider here is similar to the above except that the replacement time is determined according to the optimal replacement policy for a system under condition monitoring given by Ghasemi et al (2007). Although, this type of sequential policy has not been considered in the literature before for partially observable degrading systems, we consider it here for comparison as it is naturally a reasonable alternative for the joint policy.

We considered three combinations of cost parameters and reported the average cost per unit time on 20,000 simulated samples from the structures given earlier. We also report the average probability of a failure replacement, the average replacement interval, and the average percentage of on-time, late, and early warnings associated with all policies as given in Table 2. It can be observed from Table 2 that our model outperforms all others in terms of average cost functions. It provides worse performance in terms of the warning generation process compared to the age-based policy. The reason is that in the age-based policy, the optimal replacement time was exactly l units larger than the warning time (i.e. $T_2^* = T_1^* + l$), therefore we never have an early warning. However, age-based policy has a higher cost, higher percentage of failure replacements (except for one case), and smaller operation interval. The joint policy also outperforms the sequential policy in terms of the average cost function. Particularly, the joint policy performs much better in terms of the warning generation process. In addition, as expected, all of these policies

Table 2: Comparison with other policies

c_f	Measure	Do Nothing Policy	Age-based Policy	Sequential Policy	Joint Policy
10	g^*	2.84	1.69	1.73	1.64
	Failure Maintenance (%)	100	12.83	15.11	16.99
	Operating Interval	9.87	6.85	7.71	7.60
	On time warning (%)	0	75.83	21.19	59.33
	Late warning (%)	100	24.17	41.57	12.90
	Early warning (%)	0	0	37.24	27.77
50	g^*	6.89	2.17	2.17	2.10
	Failure Replacement (%)	100	5.63	4.15	4.25
	Operating Interval	9.87	5.94	6.25	6.13
	On time warning (%)	0	94.38	27.42	57.90
	Late warning (%)	100	5.63	30.44	13.10
	Early warning (%)	0	0	42.15	29.00
100	g^*	11.96	2.44	2.44	2.36
	Failure Replacement (%)	100	2.12	2.71	2.07
	Replacement Interval	9.87	4.98	5.81	5.42
	On time warning (%)	0	97.88	22.04	57.54
	Late warning (%)	100	2.12	22.40	9.66
	Early warning (%)	0	0	55.57	32.80

perform better than do-nothing. We should also take into account that a dynamic policy that uses online CM data for decision-making is subject to the cost of sampling and implementing a condition monitoring framework. In order to be able to make an accurate comparison between different maintenance strategies, such factors should be taken into account when the decision is one of the capital investment. In terms of the computational complexity, our model is the most expensive one due to its large state space and the many steps needed to find the optimal structure of the joint warning-replacement policy. However, since such steps are often offline, this issue is not of our concern in this work.

7.4 Investigating the Optimal Values of Static Variables

As discussed earlier, the result of this paper can be used to find the optimal values of lead time and observation interval if they are flexible. In this subsection, we show with a simple numerical example, how such an investigation can be done. Let us first assume that there is flexibility in terms of lead time (l) if we are willing to pay more to make it shorter. It is expected that by increasing the lead time, the cost of ordering decreases, but the cost of control (warning + replacement) increases. Therefore, by checking the total cost, which is the sum of the cost of control and the cost of ordering itself, one can choose the optimal lead time. An example of such investigation is shown in Figure 7 for c_s=1, c_h=0.1, c_f=50, and $c_o(l) = \{2, 1, 0.5, 0.3, 0.1, 0.1\}$ for $l \in \{1, 2, 3, 4, 5, 6\}$. For this particular setting, the optimal lead time is 4. Such analysis gives a practical tool for maintenance decision makers to analyze the effect of lead time on the cost of the system. The 2nd decision variable, which can be determined

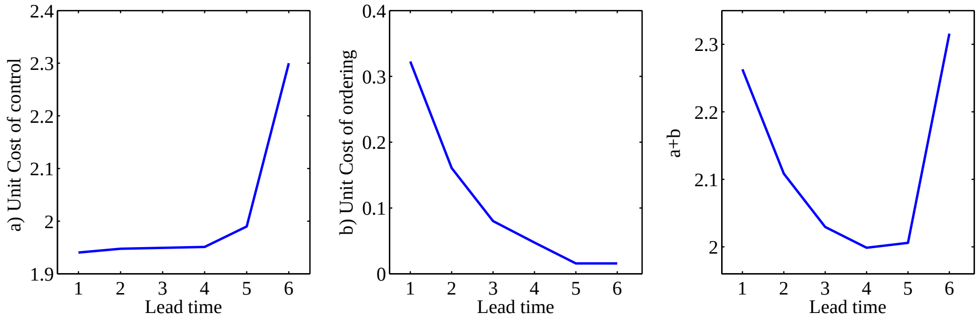

Fig. 7: Investigating the optimal lead time among 6 alternatives

offline is the observation interval. Let us assume that the fixed cost of sampling $c_{m2}(\delta)$ is 0.1 and $c_{m2}(\delta)$ is zero.

Now, as shown in Figure 8, as the observation interval becomes wider, the cost of sampling decreases, but the cost of control increases (less monitoring data is available). For this numerical setting, the optimal observation interval is 5. This type of analysis gives decision makers a tool to compare different maintenance strategies in terms of cost, likelihood of failure, replacement interval, and the percentage of late and early warning.

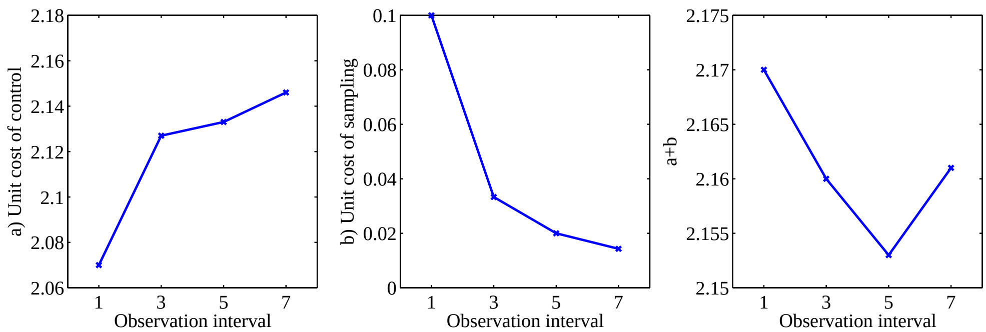

Fig. 8: Investigation for the optimal observation interval

7.5 Full Inspection Policy

One the most common maintenance strategies used in industry is the full inspection policy under which the operation is terminated for inspection. The data observed from inspection truly reflect the degradation level and therefore decision-making under full inspection is more efficient. However, inspection requires system shutdown and also is subject to other technical costs. Therefore, one should investigate the tradeoff between the cost of inspection and the cost of control in order to determine the best maintenance strategy. For $c_s = 1, c_h = 1, c_f = 50, l = 4$, we can observe from Figure 9 that inspection every 9 days has the lowest cost (unit cost of control + unit cost of inspection). Then, one can compare the best possible cost obtained from inspection with the cost of the condition monitoring framework, and then determine whether the condition monitoring framework is economically justifiable. We should point out here that, our proposed structure can perform such a comparison with using only historical condition monitoring data and without actual inspection data, which are usually very expensive to collect.

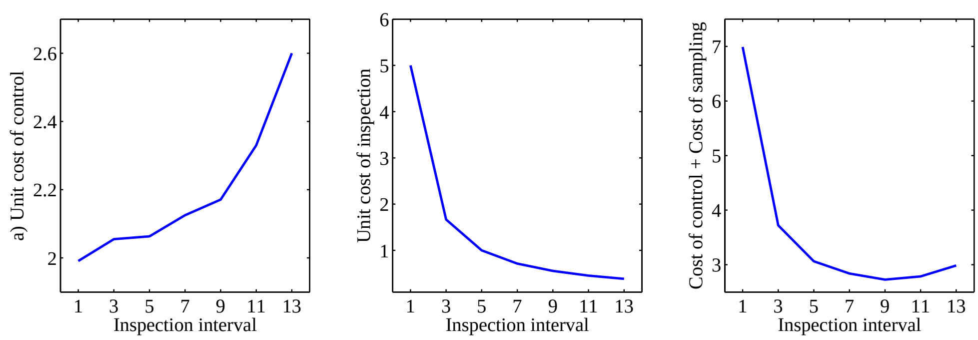

Fig. 9: Investigation for the optimal inspection interval for a Full Inspection Policy

8 Concluding Remarks and Future Research

In this paper, we developed a two-dimensional decision model, which jointly optimizes the ordering (warning) time and the maintenance time using online condition monitoring data. The objective was to utilize online condition monitoring data, which only give partial information with regards to the degradation process, for maintenance decision-making in order to minimize the expected long-run average cost per unit of operation time. A partially observable semi-Markov decision process was developed and some interesting properties of the optimal policy were discussed. In addition, simple approaches to find the optimal observation interval and optimal lead time were suggested. Future direction of this work will be on decision-making at the system level, where more than one failure modes and one components exist in the system. In addition, considering ordering size and partially observable inventory is an opportunity for further investigation.

A Proofs of Lemmas and Theorems

In this appendix, the proofs of Remark 1, Lemmas 1 - 5, Theorems 2 - 4, and Remark 2 are given.

Proof (of Remark 1) The lower bound is the expected average cost of an ideal policy under which (i) there is no failure replacement, (ii) there is no early or late warnings, and (iii) the replacement time is as close as possible to the failure time (one decision epoch before failure. Under such a policy, the system pays only for the mandatory replacement cost c_r and the expected replacement cycle is $\mathbb{E}(\zeta)$. Achieving this ideal lower bound is very hard in practice due to the stochastic nature of degradation and observation process and not guaranteed, as it requires a 100% perfect policy. The upper bound is calculated based on the fact that if an optimal policy exists, it should be at least better than (or equivalent to) the two trivial policies, (i) warning and replacement at failure and (ii) warning at time 0 and replace at failure. Therefore the cost associated with an optimal policy should be larger than or equal to the minimum of the expected costs of these two policies. In the first case, we have to pay for the cost of replacement, the cost of failure, and the cost of warning ($c_s \times l$) while the duration of a cycle is qual to $\mathbb{E}(\zeta)$. For the 2nd case, we have to pay for replacement and the cost of failure the same as case 1, however, since we issue warning at time zero, the expected cost of ordering becomes

$$\mathbb{E}\left(\mathscr{C}(\zeta,0)\right)=\int_0^l c_s(l-x)\Pr(\zeta=x)\mathrm{d}x+\int_l^\infty c_h(x-l)\Pr(\zeta=x)\mathrm{d}x.$$

Therefore, the optimal average cost is bounded, which completes the proof.

Before providing the proofs of the Theorems, we introduce some important lemmas. Using these Lemmas, we provide important results with regards to the form of decision controls, which can be used to compare two actions.

Lemma 1 *For $k<d+l$, the following term is nondecreasing in $\boldsymbol{\theta}$ and k:*

$$\int_{k\delta}^{k\delta+\delta}\mathscr{C}_l(x,d\delta)\bar{R}(\mathrm{d}x|k\delta,\boldsymbol{\theta})\;+[\mathscr{C}_l(k\delta+\delta,d\delta)]R(k\delta+\delta|k\delta,\boldsymbol{\theta})-\mathscr{C}_l(k\delta+\delta,d\delta).$$

Lemma 2 *The following holds true for all $k,d\in\mathbb{N}_0, d\le k$, and $\boldsymbol{\theta}\in\Theta$:*

$$\Big[V_\gamma(k+1,\boldsymbol{\theta},d,1)-\mathscr{C}_l(k\delta+\delta,d\delta)\Big]\ge\Big[V_\gamma(k,\boldsymbol{\theta},d,1)-\mathscr{C}_l(k\delta,d\delta)\Big].$$

Lemma 3 *If $\boldsymbol{\theta}_1\le_{st}\boldsymbol{\theta}_2$, where $\le_{st}$ means stochastically increasing (see Ohnishi et al, 1994), then the following holds true for all $k,d\in\mathbb{N}_0, d\le k$, and $\boldsymbol{\theta}_1,\boldsymbol{\theta}_2\in\Theta$:*

$$V_\gamma(k,\boldsymbol{\theta}_2,d,1)\ge V_\gamma(k,\boldsymbol{\theta}_1,d,1).$$

Lemma 4 *The value function $V_\gamma(k,\boldsymbol{\theta},\infty,1)$ is nondecreasing in $(k,\boldsymbol{\theta})$ for any $k\in\mathbb{N}_0$, and $\boldsymbol{\theta}\in\Theta$.*

Lemma 5 *The following holds true for all $k,d\in\mathbb{N}_0, d\le k$, and $\boldsymbol{\theta}\in\Theta$:*

$$\Big[V_\gamma(k+1,\boldsymbol{\theta},k,1)-\mathscr{C}_l(k\delta+\delta,k\delta)\Big]\geq\Big[V_\gamma(k,\boldsymbol{\theta},\infty,1)-\mathscr{C}_l(k\delta,k\delta)\Big].$$

Proof (of Lemma 1) This lemma also states that the given term is approximately constant in k and $\boldsymbol{\theta}$ for $k>d+l$. By simplifying $\mathscr{C}_l(k\delta,d\delta)$ and $\mathscr{C}_l(k\delta+\delta,d\delta)$ for $k<d+l$, the term in the lemma becomes

$$-c_s\delta\int_{k\delta}^{k\delta+\delta}R(\mathrm{d}x|k\delta,\boldsymbol{\theta})\;-c_skR(k\delta+\delta|k\delta,\boldsymbol{\theta})+c_s(k+1)\delta,$$

which is nondecreasing in $\boldsymbol{\theta}$ and k. We can also show that for small δ and $k>d+l$, if we assume that we issue warning at the decision epochs only, we have

$$\int_{k\delta}^{k\delta+\delta}\mathscr{C}_l(x,d\delta)\bar{R}(\mathrm{d}x|k\delta,\boldsymbol{\theta})\approx\mathscr{C}_l(k\delta+\delta,d\delta)\bar{R}(k\delta+\delta|k\delta,\boldsymbol{\theta}),\text{and thus}$$

$$\approx\mathscr{C}_l(k\delta+\delta,d\delta)\bar{R}(k\delta+\delta|k\delta,\boldsymbol{\theta})\;+[\mathscr{C}_l(k\delta+\delta,d\delta)]R(k\delta+\delta|k\delta,\boldsymbol{\theta})-\mathscr{C}_l(k\delta+\delta,d\delta)=0.$$

Proof (of Lemma 2) Since the minimum of two nondecreasing functions is nondecreasing, we prove this Lemma separately for V_γ^3 and V_γ^4. It follows from (17) that

$$V_\gamma^4(k+1,\boldsymbol{\theta},d,1)-V_\gamma^4(k,\boldsymbol{\theta},d,1)=\mathscr{C}_l((k+1)\delta,d)-\mathscr{C}_l(k\delta,d\delta),$$

which satisfies Lemma 2 if $V_\gamma(k,\boldsymbol{\theta},d,1)=V_\gamma^4(k,\boldsymbol{\theta},d,1)$. Now, it is sufficient to prove that

$$\left[V_\gamma^3(k+1,\boldsymbol{\theta},d,1)-\mathscr{C}_l(k\delta+\delta,d\delta)\right]\geq\left[V_\gamma^3(k,\boldsymbol{\theta},d,1)-\mathscr{C}_l(k\delta,d\delta)\right].$$

For a large $\hat{k}$ where $R(\hat{k}\delta+\delta|\hat{k}\delta,\boldsymbol{\theta})\approx 0$, that is when we definitely impose a preventive replacement, we have

$$V_\gamma^3(\hat{k}+1,\boldsymbol{\theta},d,1)\approx c_r+\sum_i c_f^i\theta_k^i\,\bar{R}(k\delta+\delta|k\delta,i)+V_\gamma(\boldsymbol{\pi}_0)+\mathscr{C}_l(\hat{k}\delta+\delta,d\delta)$$
$$V_\gamma^4(\hat{k}+1,\boldsymbol{\theta},d,1)\approx c_r+V_\gamma(\boldsymbol{\pi}_0)+\mathscr{C}_l(\hat{k}\delta+\delta,d\delta).$$

Therefore, $V_\gamma^4(\hat{k},\boldsymbol{\theta},d,1)<V_\gamma^3(\hat{k},\boldsymbol{\theta},d,1)$ is always true for large k, which means $V_\gamma(\hat{k},\boldsymbol{\theta},d,1)=V_\gamma^4(\hat{k},\theta,d,1)$. This implies that Lemma 1 holds true for sufficiently large k. Now, by induction hypothesis, let us assume that $[V_\gamma(k+1,\boldsymbol{\theta},d,1)-\mathscr{C}_l(k\delta+\delta,d\delta)]$ is nondecreasing for $k+1$ where $1<k+1<\hat{k}$. It follows by the definition of the value function that

$$\Big(V_\gamma(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m)-c_r-V_\gamma(\boldsymbol{\pi}_0)-\mathscr{C}_l(k\delta+\delta,d\delta)\Big)\leq 0\rightarrow$$
$$\left(\sum_{m=1}^{M}\Pr(y_{k+1}=m|\boldsymbol{\theta},u_k^2)\times V_\gamma(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),d,1)-c_r-V_\gamma(\boldsymbol{\pi}_0)-\mathscr{C}_l(k\delta+\delta,d\delta)\right)\leq 0,$$

is also nondecreasing in k. By adding $\pm\mathscr{C}_l(k\delta+\delta,d\delta)R(k\delta+\delta|k\delta,\boldsymbol{\theta}_k)$ to $V_\gamma^3(k,\boldsymbol{\theta}_k,d,1)$, we get

$$
\begin{aligned}
V^3_\gamma(k,\boldsymbol{\theta}_k,d,1) &= [c_r + V_\gamma(\boldsymbol{\pi}_0)]\,\bar{R}(k\delta+\delta|k,\boldsymbol{\theta}_k) + \sum_i c_f^i\theta_k^i\,\bar{R}(k\delta+\delta|k\delta,i) + \int\limits_{k\delta}^{k\delta+\delta} \mathscr{C}_l(x,d\delta)\bar{R}(\mathrm{d}x|k\delta,\boldsymbol{\theta}_k)\ - g(\gamma)\bar{\tau}(\delta|k,\boldsymbol{\theta}_k)\\
&+ \left(\sum_{m=1}^{M}\Pr(y_{k+1}=m|\boldsymbol{\theta}_k,1)\times V(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),d,1)\right) R(k\delta+\delta|k\delta,\boldsymbol{\theta}) \pm \mathscr{C}_l(k\delta+\delta,d\delta)R(k\delta+\delta|k\delta,\boldsymbol{\theta}_k)\\
&= (c_r + V_\gamma(\boldsymbol{\pi}_0)+)\,\bar{R}(k\delta+\delta|k,\boldsymbol{\theta}_k) + \sum_i c_f^i\theta_k^i\,\bar{R}(k\delta+\delta|k\delta,i) + \int\limits_{k\delta}^{k\delta+\delta} \mathscr{C}_l(x,d\delta)\bar{R}(\mathrm{d}x|k\delta,\boldsymbol{\theta}_k)\ - g(\gamma)\bar{\tau}(\delta|k,\boldsymbol{\theta}_k)\\
&+ \underbrace{\left(\sum_{m=1}^{M}\Pr(y_{k+1}=m|\boldsymbol{\theta}_k,1)\times V(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),d,1) - c_r - V_\gamma(\boldsymbol{\pi}_0) - \mathscr{C}_l(k\delta+\delta,d\delta)\right)}_{(*)} R(k\delta+\delta|k\delta,\boldsymbol{\theta}_k)\\
&+ \mathscr{C}_l(k\delta+\delta,d\delta)R(k\delta+\delta|k\delta,\boldsymbol{\theta}_k).
\end{aligned}
$$

As $(*)$ is nondecreasing and negative, and $R(k\delta+\delta|k\delta,\boldsymbol{\theta})$ is nonincreasing and positive in k, we can conclude that $(*)\times R(k\delta+\delta|k\delta,\boldsymbol{\theta})$ is also nondecreasing in k. Recalling Lemma 1 and the fact that R and $\bar{\tau}$ are nondecreasing in k, we conclude that $V^3_\gamma(k,\boldsymbol{\theta}_k,d,1) - \mathscr{C}_l(k\delta,d\delta)$ is nondecreasing in k. This completes the proof.

Proof (Proof of Lemma 3) It is clear that $V^4_\gamma(k,\boldsymbol{\theta},d,1)$ is constant wrt $\boldsymbol{\theta}$ and therefore is nondecreasing in $\boldsymbol{\theta}$. The rest of the proof is to show that $V^3_\gamma(k,\boldsymbol{\theta},d,1)$ is nondecreasing in $\boldsymbol{\theta}$. Following the same steps as we did in Lemma 2 , we get

$$
V^3_\gamma(k,\boldsymbol{\theta}_2,d,1) - V^3_\gamma(k,\boldsymbol{\theta}_1,d,1) \geq \left[\int_{k\delta}^{k\delta+\delta} \mathscr{C}_l(x,d\delta)\bar{R}(\mathrm{d}x|k\delta,\boldsymbol{\theta})\ + [\mathscr{C}_l(k\delta+\delta,d\delta)]R(k\delta+\delta|k\delta,\boldsymbol{\theta})\right]\Bigg|_{\boldsymbol{\theta}_1}^{\boldsymbol{\theta}_2} \geq 0.
$$

This completes the proof. This Lemma implies that if the decision process starts from a state that is stochastically more deteriorated in the sense of the stochastic ordering, it will be incurred a higher or equal cost in future.

Proof (of Lemma 4) We first show the proof for k. The proof simply includes verifying that all elements in the right-hand side of (12) $(V^0_\gamma(k,\boldsymbol{\theta},+\infty,1) - V^2_\gamma(k,\boldsymbol{\theta},+\infty,1))$ are nondecreasing functions in k given the fact that the minimum of nondecreasing functions is also nondecreasing. From (13)-(15), we can show that for a sufficiently large $\hat{k}$ (where $R(k\delta+\delta|k\delta,\boldsymbol{\theta})\approx 0$), we have

$$
V^i_\gamma(\hat{k},\boldsymbol{\theta},\infty,1) \approx \begin{cases} c_r + \sum_i c_f^i\theta^i\,\bar{R}(k\delta+\delta|k\delta,i) + V_\gamma(\boldsymbol{\pi}_0) + \mathscr{C}_l(\hat{k}\delta,\hat{k}\delta), & i=0,1\\ c_r + V_\gamma(\boldsymbol{\pi}_0) + \mathscr{C}_l(\hat{k}\delta,\hat{k}\delta), & i=2 \end{cases}.
$$

Therefore, the statement $V^2_\gamma(\hat{k},\boldsymbol{\theta},+\infty,1) < \min\{V^0_\gamma(\hat{k},\boldsymbol{\theta},+\infty,1), V^1_\gamma(k,\boldsymbol{\theta},+\infty,1)\}$ is always true regardless of $\boldsymbol{\theta}$. This implies that when the age of the device is very high in the sense that it is very likely to fail during the next observation interval (given that the warning has not been generated), the best decision is to *replace immediately.* It is also obvious that $[c_r + c_s l + V_\gamma(\boldsymbol{\pi}_0)]$ is constant in k and therefore $V^2_\gamma(k,\boldsymbol{\theta},+\infty,1)$ is nondecreasing in k. Therefore $V_\gamma(k,\boldsymbol{\theta},\infty,1)$ is constant and nondecreasing for sufficiently large k. The rest of the proof is to show that $V^0_\gamma(k,\boldsymbol{\theta},+\infty,1)$ and $V^1_\gamma(k,\boldsymbol{\theta},+\infty,1)$ are both nondecreasing in k for all $k<\hat{k}$. Let us first assume by induction hypothesis that $V_\gamma(k+1,\boldsymbol{\theta},+\infty,1)$ is nondecreasing in k. It follows by definition that

$$
\left[\sum_{m=1}^{M}\Pr(y_{k+1}=m|\boldsymbol{\theta},1)\times V_\gamma(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),\infty,1)\right] - [c_r + c_s l + V_\gamma(\boldsymbol{\pi}_0)] \leq 0.
$$

Since $R(k\delta+\delta|k\delta,\boldsymbol{\theta})$ and $V_\gamma(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),\infty,1)$ are nondecreasing in k, then

$$
(**) = \left[\sum_{m=1}^{M}\Pr(y_{k+1}=m|\boldsymbol{\theta},1)\times V_\gamma(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),\infty,1) - [c_r + c_s l + V_\gamma(\boldsymbol{\pi}_0)]\right] R(k\delta+\delta|k\delta,\boldsymbol{\theta}),
$$

is also nondecreasing in k. Also, from (13), we know that

$$V_\gamma^0(k,\boldsymbol{\theta},\infty,1) = (**) + c_r + V_\gamma(\boldsymbol{\pi}_0) + \sum_i c_f^i \theta^i \,\bar{R}(k\delta+\delta|k\delta,i) - g(\gamma)\bar{\tau}(\delta|k,\boldsymbol{\theta},1) + \mathscr{C}_l(k\delta,k\delta).$$

Since all of the above elements are nondecreasing in k, therefore $V_\gamma^0(k,\boldsymbol{\theta},+\infty,1)$ is nondecreasing in k. Implementing the same procedures as we did above and considering Lemma 2, which states that $V_\gamma(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),k,1) - \mathscr{C}_l(k\delta+\delta,k\delta)$ is nondecreasing in k, we have

$$(***) = \left[\sum_{m=1}^{M} \Pr(y_{k+1}=m|\boldsymbol{\theta},1) \times V_\gamma(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),k,1) - [c_r + \mathscr{C}_l(k\delta+\delta,k\delta) + V_\gamma(\boldsymbol{\pi}_0)]\right] \le 0,$$

which is also nondecreasing in k. From (13), we have

$$V_\gamma^1(k,\boldsymbol{\theta},+\infty,1) = (***)R(k\delta+\delta|k\delta,\boldsymbol{\theta}_k) + \sum_i c_f^i \theta^i \,\bar{R}(k\delta+\delta|k\delta,i) - g(\gamma)\bar{\tau}(\delta|k,\boldsymbol{\theta})$$
$$+ \mathscr{C}_l(k\delta+\delta,k\delta)R(k\delta+\delta|k\delta,\boldsymbol{\theta}) + \int_{k\delta}^{k\delta+\delta} \mathscr{C}_l(x,k\delta)\bar{R}(\mathrm{d}x|k\delta,\boldsymbol{\theta}_k) + c_r + V(\boldsymbol{\pi}_0).$$

As all of the above terms are nondecreasing in k, then $V_\gamma^1(k,\boldsymbol{\theta},+\infty,1)$ is also nondecreasing in k. Considering that $V_\gamma^i(k,\boldsymbol{\theta},+\infty,1)$ is a nondecreasing function in k for $i \in \{0,1,2\}$, we can conclude that $V_\gamma(k,\boldsymbol{\theta},\infty,1)$ is also nondecreasing in k. This completes the proof for k. To prove that $V_\gamma(k,\boldsymbol{\theta},\infty,1)$ is nondecreasing in $\boldsymbol{\theta}$, we can implement the same steps as we did for the case of k, except that we use Lemma 3 to show that $(***)$ and $(***)R(k\delta+\delta|k\delta,\boldsymbol{\theta}_k)$ are nondecreasing in $\boldsymbol{\theta}$. We have removed the rest of the proof for its simplicity. This completes the proof.

Proof (of Lemma 5) From (14) and (17), we have $V_\gamma^1(k,\boldsymbol{\theta}_k,\infty,1) = V_\gamma^3(k,\boldsymbol{\theta}_k,k,1)$. From the results given in the proof of Lemma 2, we know that

$$V_\gamma(k,\boldsymbol{\theta}_k,\infty,1) \le V_\gamma^3(k,\boldsymbol{\theta}_k,k,1) \le V_\gamma^3(k+1,\boldsymbol{\theta}_k,k,1) - [\mathscr{C}_l(k\delta+\delta,k\delta) - \mathscr{C}_l(k\delta,k\delta)]\,.$$

Recall from (15) that

$V_\gamma(k,\boldsymbol{\theta}_k,\infty,1) \le c_r + \mathscr{C}_l(k\delta,k\delta) + V_\gamma(\boldsymbol{\pi}_0) = V_\gamma^4(k1,\boldsymbol{\theta}_k,k,1)$. By adding $\pm[\mathscr{C}_l(k\delta+\delta,k\delta) - \mathscr{C}_l(k\delta,k\delta)]$ to the right-hand side of the above inequality, we have

$$V_\gamma(k,\boldsymbol{\theta}_k,\infty,1) \le V_\gamma^4(k+1,\boldsymbol{\theta}_k,k,1) - [\mathscr{C}_l(k\delta+\delta,k\delta) - \mathscr{C}_l(k\delta,k\delta)]\,.$$

Recalling that $V_\gamma(k+1,\boldsymbol{\theta}_k,k,1) = \min\{V_\gamma^3(k+1,\boldsymbol{\theta}_k,k,1), V_\gamma^4(k+1,\boldsymbol{\theta}_k,k,1)\}$, we get

$$[V_\gamma(k+1,\boldsymbol{\theta},k,1) - V_\gamma(k,\boldsymbol{\theta},\infty,1)] \ge [\mathscr{C}_l(k\delta+\delta,k\delta) - \mathscr{C}_l(k\delta,k\delta)]\,,$$

which completes the proof.

Proof (of Theorem 1) From (14)-(15), we have

$$V_\gamma^1(\boldsymbol{\pi}_k) - V_\gamma^2(\boldsymbol{\pi}_k) = \sum_i c_f^i \theta_k^i \,\bar{R}(k\delta+\delta|k\delta,i) - \mathscr{C}_l(k\delta,k\delta)\bar{R}(k\delta+\delta|k\delta,\boldsymbol{\theta}_k) + \int_{k\delta}^{k\delta+\delta} \mathscr{C}_l(x,k\delta)\bar{R}(\mathrm{d}x|k\delta,\boldsymbol{\theta}_k) - g(\gamma)\bar{\tau}(\delta|k,\boldsymbol{\theta}_k)$$
$$+ \left[\sum_{m=1}^{M} \Pr(y_{k+1}=m|\boldsymbol{\theta}_k,1) \times V_\gamma(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),k,1) - V_\gamma^2(k,\boldsymbol{\theta}_k,\infty,1)\right] R(k\delta+\delta|k\delta,\boldsymbol{\theta}_k).$$

Replacing $V_\gamma^2(k+1,\boldsymbol{\theta}_{k+1},\infty,1)$ by $V_\gamma^4(k+1,\boldsymbol{\theta}_{k+1},k,1) + c_s$ from (19) and adding $\mathscr{C}_l(k\delta+\delta,k\delta) - \mathscr{C}_l(k\delta,k\delta)$, we get

$$V_\gamma^1(\boldsymbol{\pi}_k) - V_\gamma^2(\boldsymbol{\pi}_k) = \sum_i c_f^i \theta_k^i \,\bar{R}(k\delta+\delta|k\delta,i) - \mathscr{C}_l(k\delta+\delta,k\delta)\bar{R}(k\delta+\delta|k\delta,\boldsymbol{\theta}_k) - g(\gamma)\bar{\tau}(\delta|k,\boldsymbol{\theta}_k) - c_s + \int_{k\delta}^{k\delta+\delta} \mathscr{C}_l(x,k\delta)\bar{R}(\mathrm{d}x|k\delta,\boldsymbol{\theta}_k) \tag{31}$$

$$+ \left[\sum_{m=1}^{M} \Pr(y_{k+1}=m|\boldsymbol{\theta}_k,1) \times V_\gamma(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),k,1) - V_\gamma^4(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),k,1)\right] R(k\delta+\delta|k\delta,\boldsymbol{\theta}_k). \tag{32}$$

Recall from (16) that $V_\gamma(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),k,1) \leq V_\gamma^4(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),k,1)$, and thus

$$\begin{aligned}\sum_{m=1}^{M} \Pr(y_{k+1}=m|\boldsymbol{\theta}_k,1)V_\gamma(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),k,1) &\leq \sum_{m=1}^{M} \Pr(y_{k+1}=m|\boldsymbol{\theta}_k,1)V_\gamma^4(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),k,1)\\ &\leq V_\gamma^4(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),k,1),\end{aligned}$$

which follows that the last term in (31) is non-positive. Therefore if $\Phi_\gamma^1(\boldsymbol{\pi}_k)$ is non-positive, then $V_\gamma^1(\boldsymbol{\pi}_k) \leq V_\gamma^2(\boldsymbol{\pi}_k)$, that is *warning immediately and no replacement* is better than *immediate warning and immediate replacement*. Now we show that if $\Phi_\gamma^1(k,\boldsymbol{\theta}_k,\infty,1) > 0$, then $V_\gamma(\boldsymbol{\pi}_k) = V_\gamma^1(\boldsymbol{\pi}_k)$ cannot be true. Let us assume that $\Phi_\gamma^1(\boldsymbol{\pi}_k) > 0$ and $V_\gamma(\boldsymbol{\pi}_k) = V_\gamma^1(\boldsymbol{\pi}_k)$. Then we have

$$\begin{aligned}V_\gamma(k+1,\boldsymbol{\theta}_k,k,1) - V_\gamma(\boldsymbol{\pi}_k) &= V(k+1,\boldsymbol{\theta}_k,k,1) - V_\gamma^1(\boldsymbol{\pi}_k)\\ &= -\Phi_\gamma^1(\boldsymbol{\pi}_k) + \Big[V_\gamma(k+1,\boldsymbol{\theta}_k,k,1) - c_r - \mathscr{C}_l(k\delta+\delta,k\delta) - V_\gamma(\boldsymbol{\pi}_0)\Big](1-R(k\delta+\delta|k\delta,\boldsymbol{\theta}_k))\\ &+ \left[V_\gamma(k+1,\boldsymbol{\theta}_k,k,1) - \sum_{m=1}^{M}\Pr(y_{k+1}=m|\boldsymbol{\theta}_k,1)\times V_\gamma(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),k,1)\right]R(k\delta+\delta|k\delta,\boldsymbol{\theta}_k) - c_s.\end{aligned}$$

From Lemma 5, we know that the left-hand side is greater than $\mathscr{C}_l(k\delta+\delta,k\delta) - \mathscr{C}_l(k\delta,k\delta)$. Therefore, since $-\Phi_\gamma^1(\boldsymbol{\pi}_k) < 0$, the second term on the right-hand side (see Equation (16)), and the third term on the right-hand side (see Lemma 3) are less than zero, then the right-hand side of the above equation is less than $\mathscr{C}_l(k\delta+\delta,k\delta) - \mathscr{C}_l(k\delta,k\delta) = -c_s$. Thus, the above equation cannot be true as the sign of the two sides of the equation does not match. This means that if $\Phi_\gamma^1(\boldsymbol{\pi}_k) > 0$, then $V_\gamma(\boldsymbol{\pi}_k) < V_\gamma^1(\boldsymbol{\pi}_k)$ cannot be true.

Proof (of Theorem 2) From (13)-(14), we have

$$\begin{aligned}V_\gamma^0(k,\boldsymbol{\theta}_k,\infty,1) - V_\gamma^2(k,\boldsymbol{\theta}_k,\infty,1) &= \underbrace{\sum_i c_f^i\theta_k^i\,\bar{R}(k\delta+\delta|k\delta,i) - g(\gamma)\bar{\tau}(\delta|k,\boldsymbol{\theta}_k)}_{\Phi_\gamma^2(t,\boldsymbol{\theta}_k,\infty,1)}\\ &+ \left[\sum_m \Pr(y_{k+1}=m|\boldsymbol{\theta}_k,1)\times V_\gamma(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),\infty,1) - V_\gamma^2(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),\infty,1)\right]R(k\delta+\delta|k\delta,\boldsymbol{\theta}_k).\end{aligned}$$

As the second term of the above is nonpositive, it is obvious that if $\Phi_\gamma^2(k,\boldsymbol{\theta}_k,\infty,1) \leq 0$, then $V_\gamma^0(k,\boldsymbol{\theta}_k,\infty,1) \leq V_\gamma^2(k,\boldsymbol{\theta}_k,\infty,1)$, that is, the *do nothing* is a better action. Now, we can show that if $\Phi_\gamma^2(k,\boldsymbol{\theta}_k,\infty,1) > 0$, then $V_\gamma(k,\boldsymbol{\theta}_k,\infty,1) = V_\gamma^0(k,\boldsymbol{\theta}_k,\infty,1)$ cannot be true. Thus, if $\boldsymbol{\pi}_k = (k,\boldsymbol{\theta}_k,+\infty,1)$ and $\boldsymbol{\pi}_{k+1} = (k+1,\boldsymbol{\theta}_k,+\infty,1)$, then

$$
\begin{aligned}
V_\gamma(\boldsymbol{\pi}_{k+1}) - V_\gamma(\boldsymbol{\pi}_k) = V_\gamma(\boldsymbol{\pi}_{k+1}) - V_\gamma^0(\boldsymbol{\pi}_k) > 0 \\
&= V_\gamma(\boldsymbol{\pi}_{k+1}) - \Phi_\gamma^2(\boldsymbol{\pi}_k) - [c_r + \mathscr{C}_l(k\delta, k\delta) + V_\gamma(\boldsymbol{\pi}_0)](1 - R(k\delta+\delta|k\delta, \boldsymbol{\theta}_k)) \\
&- \left[\sum_{m=1}^{M} \Pr(y_{k+1} = m|\boldsymbol{\theta}_k, 1) \times V_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), \infty, 1)\right] R(k\delta+\delta|k\delta, \boldsymbol{\theta}_k) \times V_\gamma(k+1, \boldsymbol{\theta}_k, +\infty, 1) \\
&= -\Phi_\gamma^2(\boldsymbol{\pi}_k) + [V_\gamma(\boldsymbol{\pi}_{k+1}) - c_r - \mathscr{C}_l(k\delta, k\delta) - V_\gamma(\boldsymbol{\pi}_0)](1 - R(k\delta+\delta|k\delta, \boldsymbol{\theta}_k)) \\
&+ \left[V_\gamma(\boldsymbol{\pi}_k) - \sum_{m=1}^{M} \Pr(y_{k+1} = m|\boldsymbol{\theta}_k, 1) \times V_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), \infty, 1)\right] R(k\delta+\delta|k\delta, \boldsymbol{\theta}_k).
\end{aligned}
$$

As $-\Phi_\gamma^2(\boldsymbol{\pi}_k) < 0$, and the second and the third terms are also nonpositive, we can conclude that the above equation cannot be true. Therefore, if $\Phi_\gamma^2(\boldsymbol{\pi}_k) > 0$, then $V_\gamma(\boldsymbol{\pi}_k) = V_\gamma^0(\boldsymbol{\pi}_k)$ cannot be true.

Proof (of Theorem 3) Using (13)-(14), we have

$$
\begin{aligned}
V_\gamma^0(\boldsymbol{\pi}_k) - V_\gamma^1(\boldsymbol{\pi}_k) = \mathscr{C}_l(k\delta, k\delta)(1 - R(k\delta+\delta|k\delta, \boldsymbol{\theta}_k)) - \int_{k\delta}^{k\delta+\delta} \mathscr{C}_l(x, k\delta)\bar{R}(\mathrm{d}x|k\delta, \boldsymbol{\theta}_k) \\
+ \left[\sum_{m=1}^{M} \Pr(y_{k+1} = m|\boldsymbol{\theta}_k, 1) \times [V_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), \infty, 1) - V_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), k, 1)]\right] R(k\delta+\delta|k\delta, \boldsymbol{\theta}_k).
\end{aligned}
\tag{33}
$$

The comparison between the value function of these two options requires evaluation of the last term of the above equation, which can be rewritten as

$$
\left[\mathbb{E}^m \left[V_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), \infty, 1)\right] - \mathbb{E}^m \left[V_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), k, 1)\right]\right] R(k\delta+\delta|k\delta, \boldsymbol{\theta}_k),
$$

where $\mathbb{E}^m$ is the expectation operation wrt to m, given that the replacement control is to *do nothing*. To calculate the above term, we use the property of the value function as

$$
\begin{aligned}
V_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), \infty, 1, \gamma) - V_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), k, 1) = \sum_i c_f^i \times \Big(\bar{Q}_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), \infty, 1|i) - \bar{Q}_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), k, 1|i)\Big) \\
+ \Big(\bar{C}_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), \infty, 1) - \bar{C}_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), k, 1)\Big) \\
- g(\gamma)\Big(\bar{W}_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), \infty, 1) - \bar{W}_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), k, 1)\Big).
\end{aligned}
$$

where $\bar{W}_\gamma(\boldsymbol{\pi}_k)$, $\bar{Q}_\gamma(\boldsymbol{\pi}_k|i)$, and $\bar{C}_\gamma(\boldsymbol{\pi}_k)$ (see §5.4.1) are the expected remaining time to replacement, the expected probability of failure at state i, and the expected cost of the warning process, given the state $\boldsymbol{\pi}_k$ and a policy γ with cost $g(\gamma)$. Now, the statement given in Theorem 3 follows.

Proof (of Theorem 4) As shown earlier, when the warning has already been issued, there are only two maintenance options to compare at the policy improvement step. Thus, if we assume $\boldsymbol{\pi}_k = (k, \boldsymbol{\theta}_k, d, 1)$, we get

$$
\begin{aligned}
V_\gamma^3(\boldsymbol{\pi}_k) - V_\gamma^4(\boldsymbol{\pi}_k) = [c_r + V_\gamma(\boldsymbol{\pi}_0)]\,(\bar{R}(k\delta+\delta|k\delta, \boldsymbol{\theta}_k)) + \sum_i c_f^i \theta_k^i\, \bar{R}(k\delta+\delta|k\delta, i) + \int_{k\delta}^{k\delta+\delta} \mathscr{C}_l(x, d\delta)\bar{R}(\mathrm{d}x|k\delta, \boldsymbol{\theta}_k) \\
- g(\gamma)\bar{\tau}(\delta|k, \boldsymbol{\theta}_k) + \left(\sum_{m=1}^{M} \Pr(y_{k+1} = m|\boldsymbol{\theta}_k, 1) \times V_\gamma(k+1, \hat{\boldsymbol{\theta}}_{k+1}(m), d, 1)\right) R(k\delta+\delta|k\delta, \boldsymbol{\theta}_k) \\
- [c_r + \mathscr{C}_l(k\delta, d\delta) + V_\gamma(\boldsymbol{\pi}_0)]\,.
\end{aligned}
$$

By adding $\pm\mathscr{C}_l(k\delta+\delta, d\delta)R(k\delta+\delta|k\delta,\boldsymbol{\theta}_k)$ to the above, we get

$$V_\gamma^3(\boldsymbol{\pi}_k) - V_\gamma^4(\boldsymbol{\pi}_k)$$

$$= \underbrace{\sum_i c_f^i \theta_k^i \, \bar{R}(k\delta+\delta|k\delta,i) - g(\gamma)\bar{\tau}(\delta|k,\boldsymbol{\theta}_k) + [\mathscr{C}_l(k\delta+\delta,d\delta)R(k\delta+\delta|k\delta,\boldsymbol{\theta}_k) - \mathscr{C}_l(k\delta,d\delta)] + \int_{k\delta}^{k\delta+\delta} \mathscr{C}_l(x,d\delta)\bar{R}(\mathrm{d}x|k\delta,\boldsymbol{\theta}_k)}_{\Phi_\gamma^4(k,\boldsymbol{\theta}_k,d,1)}$$

$$+ \underbrace{\left[\sum_{m=1}^{M} \Pr(y_{k+1}=m|\boldsymbol{\theta}_k,1) \times V_\gamma(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),d,1) - V_\gamma^4(k+1,\boldsymbol{\theta}_{k+1},d,1)\right] R(k\delta+\delta|k\delta,\boldsymbol{\theta}_k)}_{(****)}.$$

It follows from (16) that $V_\gamma(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),\infty,1) \le V_\gamma^4(k+1,\boldsymbol{\theta}_{k+1},d,1)$, and thus

$$\sum_{m=1}^{M} \Pr(y_{k+1}=m|\boldsymbol{\theta}_k,1) \times V_\gamma(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),\infty,1) \le \sum_{m=1}^{M} \Pr(y_{k+1}=m|\boldsymbol{\theta}_k,1) V_\gamma^4(k+1,\boldsymbol{\theta}_{k+1},d,1),$$

which means $(****) < 0$. Therefore, if $\Phi_\gamma^4(\boldsymbol{\pi}_k) \le 0$, then $V_\gamma^3(\boldsymbol{\pi}_k) - V_\gamma^4(\boldsymbol{\pi}_k) < 0$, that is the optimal policy is *do nothing.* The rest of the proof is devoted to show that if $\Phi_\gamma^4(\boldsymbol{\pi}_k) > 0$, then $V_\gamma^3(\boldsymbol{\pi}_k) - V_\gamma^4(\boldsymbol{\pi}_k) < 0$ cannot be true and therefore $V_\gamma^3(\boldsymbol{\pi}_k) > V_\gamma^4(\boldsymbol{\pi}_k)$, that is the optimal policy is to *replace immediately.* It follows from Lemma 2 that $V_\gamma(k+1,\boldsymbol{\theta},d,1) - V(\boldsymbol{\pi}_k) > [\mathscr{C}_l(k\delta+\delta,d\delta) - \mathscr{C}_l(k\delta,d\delta)]$ and then

$$V_\gamma(k+1,\boldsymbol{\theta},d,1) - V_\gamma(\boldsymbol{\pi}_k) = V_\gamma(k+1,\boldsymbol{\theta},d,1) - V_\gamma^3(\boldsymbol{\pi}_k)$$

$$= V_\gamma(k+1,\boldsymbol{\theta},k,1) - V_\gamma^3(\boldsymbol{\pi}_k) + [V_\gamma(k+1,\boldsymbol{\theta},d,1)R(k\delta+\delta|k\delta,\boldsymbol{\theta}) - V_\gamma(k+1,\boldsymbol{\theta},d,1)R(k\delta+\delta|k\delta,\boldsymbol{\theta})]$$

$$= -\Phi_\gamma^4(k,\boldsymbol{\theta},d,1) + [\mathscr{C}_l(k\delta+\delta,d\delta)R(k\delta+\delta|k\delta,\boldsymbol{\theta}) - \mathscr{C}_l(k\delta,d\delta)]$$

$$+ [V_\gamma(k+1,\boldsymbol{\theta},d,1) - c_r - V_\gamma(\boldsymbol{\pi}_0)](1 - R(k\delta+\delta|k\delta,\boldsymbol{\theta}))$$

$$+ V_\gamma(k+1,\boldsymbol{\theta},d,1)R(\delta|k,\theta) - \left[\sum_{m=1}^{M} \Pr(y_{k+1}=m|\boldsymbol{\theta},1) \times V_\gamma(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),d,1)\right] R(k\delta+\delta|k\delta,\boldsymbol{\theta}_k).$$

By adding $\mathscr{C}_l(k\delta+\delta,d\delta) - \mathscr{C}_l(k\delta+\delta,d\delta)$ to the above, we get

$$V_\gamma(k+1,\boldsymbol{\theta},d,1) - V_\gamma(\boldsymbol{\pi}_k) = \underbrace{-\Phi^4(k,\boldsymbol{\theta},d,1)}_{<0} + [\mathscr{C}_l(k\delta+\delta,d\delta) - \mathscr{C}_l(k\delta,d\delta)]$$

$$+ [V_\gamma(k+1,\boldsymbol{\theta},d,1) - c_r - \mathscr{C}_l(k\delta+\delta,d\delta) - V_\gamma(\boldsymbol{\pi}_0)](1 - R(k\delta+\delta|k\delta,\boldsymbol{\theta}))$$

$$+ \left[V_\gamma(k+1,\boldsymbol{\theta},d,1) - \sum_{m=1}^{M} \Pr(y_{k+1}=m|\boldsymbol{\theta},1) \times V_\gamma(k+1,\hat{\boldsymbol{\theta}}_{k+1}(m),d,1)\right] R(k\delta+\delta|k\delta,\boldsymbol{\theta}).$$

Therefore, as the last two terms of the above are nonpositive, if $\Phi_\gamma^4(\boldsymbol{\pi}_k) > 0$, we have sharp inequality, which is a contradiction. From the above we can conclude that $V_\gamma^3(\boldsymbol{\pi}_k) > V_\gamma^4(\boldsymbol{\pi}_k)$, that is the optimal decision is to *replace immediately.* This completes the proof.

Proof (of Remark 2) By taking the difference between Φ_γ^1 and Φ_γ^2 , we get

$$\Phi_\gamma^1(k,\boldsymbol{\theta}_k,\infty,1) - \Phi_\gamma^2(k,\boldsymbol{\theta}_k,\infty,1) = \mathscr{C}_l(k\delta+\delta,k\delta) - \mathscr{C}_l(k\delta,k\delta) \le 0,$$

which completes the proof.

Algorithm 1: The Backward Recursive Process to find $\bar{Q}$, $\bar{W}$, and $\bar{C}$ for a given policy γ

Step 1. Using the unconditional reliability function, find T_{max} as a time point at which the reliability is approximately 0. This makes Φ_1-Φ_4 greater than zero with probability 1. Set $k := \lceil T_{max}/\delta \rceil$, and move to Step 2.

Step 2. For $d := 0 : k$, and each $\boldsymbol{\theta}_k \in \tilde{\Theta}_\sigma$, update $\boldsymbol{\pi_k}$. If $\gamma_2(\boldsymbol{\pi}_k) = 1$, calculate $\bar{W}_\gamma(\boldsymbol{\pi}_k)$, $\bar{Q}_\gamma(\boldsymbol{\pi}_k, i), \forall i$, and $\bar{C}_\gamma(\boldsymbol{\pi}_k)$, and if $\gamma_2(\boldsymbol{\pi}_k) = 0$, calculate $\bar{C}_\gamma(\boldsymbol{\pi}_k)$ only.

Step 3. Set $k := k - 1$. If $k \geq 0$, move back to Step 2, otherwise go to Step 4.

Step 4. Let $\lceil k = T_{max}/\delta \rceil$. For $d = \infty$ and $\boldsymbol{\theta}_k \in \tilde{\Theta}_\epsilon$, update $\boldsymbol{\pi_k}$. If $\gamma_2(\boldsymbol{\pi}_k) = 1$, calculate $\bar{W}_\gamma(\boldsymbol{\pi}_k)$, $\bar{Q}_\gamma(\boldsymbol{\pi}_k, i)$, and $\bar{C}_\gamma(\boldsymbol{\pi}_k)$, and if $\gamma_2(\boldsymbol{\pi}_k) = 0$, calculate $\bar{C}_\gamma(\boldsymbol{\pi}_k)$ only.

Step 5. Set $k := k - 1$. If $k \geq 0$, move to Step 4, otherwise terminate the algorithm and compute $g(\gamma)$ from (9).

References

Ahmad R, Kamaruddin S (2012) A review of condition-based maintenance decision-making. European Journal of Industrial Engineering 6(5):519–541

Armstrong M, Atkins D (1996) Joint optimization of maintenance and inventory policies for a simple system. IIE Transactions (Institute of Industrial Engineers) 28(5):415–424

Armstrong M, Atkins D (1998) A note on joint optimization of maintenance and inventory. IIE Transactions (Institute of Industrial Engineers) 30(2):143–149

Aven T, Bergman B (1986) Optimal replacement times - a general set-up. Journal of Applied Probability 23(2):432–442

Banjevic D, Jardine AKS (2006) Calculation of reliability function and remaining useful life for a markov failure time process. IMA Journal of Management Mathematics 17(2):115–130

Banjevic D, Jardine A, Makis V, Ennis M (2001) A control-limit policy and software for condition-based maintenance optimization. INFOR 39(1):32–50

Bertsekas DP, Tsitsiklis JN (1996) Neuro-Dynamic Programming, 1st edn. Athena Scientific

Brezavscek A, Hudoklin A (2003) Joint optimization of block-replacement and periodic-review spare-provisioning policy. IEEE Transactions on Reliability 52(1):112–117

Chen A, Wu G (2007) Real-time health prognosis and dynamic preventive maintenance policy for equipment under aging markovian deterioration. International Journal of Production Research 45(15):3351–3379

Dong M, He D (2007) Hidden semi-markov model-based methodology for multi-sensor equipment health diagnosis and prognosis. European Journal of Operational Research 178(3):858–878

Elwany A, Gebraeel N (2008) Sensor-driven prognostic models for equipment replacement and spare parts inventory. IIE Transactions (Institute of Industrial Engineers) 40(7):629–639

Ghasemi A, Yacout S, Ouali M (2007) Optimal condition based maintenance with imperfect information and the proportional hazards model. International Journal of Production Research 45(4):989–1012

Ghasemi A, Yacout S, Ouali MS (2010) Parameter estimation methods for condition-based maintenance with indirect observations. IEEE Transactions on Reliability 59(2):426–439

Godoy D, Pascual R, Knights P (2013) Critical spare parts ordering decisions using conditional reliability and stochastic lead time. Reliability Engineering and System Safety 119:199–206

Hsu BM, Shu MH (2010) Reliability assessment and replacement for machine tools under wear deterioration. International Journal of Advanced Manufacturing Technology 48(1-4):355–365

Ivy J, Pollock S (2005) Marginally monotonic maintenance policies for a multi-state deteriorating machine with probabilistic monitoring, and silent failures. IEEE Transactions on Reliability 54(3):489–497

Jardine A, Lin D, Banjevic D (2006) A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical Systems and Signal Processing 20(7):1483–1510

Kabir A, Farrash S (1996) Simulation of an integrated age replacement and spare provisioning policy using slam. Reliability Engineering and System Safety 52(2):129–138

Kim M, Makis V (2013) Joint optimization of sampling and control of partially observable failing systems. Operations Research 61(3):777–790

Kurt M, Kharoufeh J (2010) Monotone optimal replacement policies for a markovian deteriorating system in a controllable environment. Operations Research Letters 38(4):273–279

Lam C, Yeh R (1994) Optimal replacement policies for multistate deteriorating systems. Naval Research Logistics 41(3):303–315

Louit D, Pascual R, Banjevic D, Jardine A (2011) Condition-based spares ordering for critical components. Mechanical Systems and Signal Processing 25(5):1837–1848

Maillart L (2006) Maintenance policies for systems with condition monitoring and obvious failures. IIE Transactions (Institute of Industrial Engineers) 38(6):463–475

Makis V, Jardine A (1992) Optimal replacement in the proportional hazards model. INFOR 30(1):172183

Ohnishi M, Morioka T, Ibaraki T (1994) Optimal minimal-repair and replacement problem of discrete-time markovian deterioration system under incomplete state information. Computers and Industrial Engineering 27(1-4):409–412

Panagiotidou S (2014) Joint optimization of spare parts ordering and maintenance policies for multiple identical items subject to silent failures. European Journal of Operational Research 235(1):300–314

Peng Y, Dong M, Zuo M (2010) Current status of machine prognostics in condition-based maintenance: A review. International Journal of Advanced Manufacturing Technology 50(1-4):297–313

Qian X, Wu Y (2014) Condition based maintenance optimization for the hydro generating unit with dynamic economic dependence. International Journal of Control and Automation 7(3):317–326

Rosenfield D (1976) Markovian deterioration with uncertain information - a more general model. Naval Research Logistics 23(3):389–405

Sun J, Zuo H, Wang W, Pecht M (2012) Application of a state space modeling technique to system prognostics based on a health index for condition-based maintenance. Mechanical Systems and Signal Processing 28:585–596

Teng H, Zhao J, Jia X, Jia Y, Zhang X, Cai L (2011) Experimental study on gearbox prognosis using total life vibration analysis

Van Horenbeek A, Bur J, Cattrysse D, Pintelon L, Vansteenwegen P (2013) Joint maintenance and inventory optimization systems: A review. International Journal of Production Economics 143(2):499–508

Wang L, Chu J, Mao W (2008a) A condition-based order-replacement policy for a single-unit system. Applied Mathematical Modelling 32(11):2274–2289

Wang L, Chu J, Mao W (2008b) An optimum condition-based replacement and spare provisioning policy based on markov chains. Journal of Quality in Maintenance Engineering 14(4):387–401

Wang L, Chu J, Mao W (2009) A condition-based replacement and spare provisioning policy for deteriorating systems with uncertain deterioration to failure. European Journal of Operational Research 194(1):184–205

Wang Z, Hu C, Wang W, Kong X, Zhang W (2015) A prognostics-based spare part ordering and system replacement policy for a deteriorating system subjected to a random lead time. International Journal of Production Research 53(15):4511–4527

Wu X, Ryan S (2014) Joint optimization of asset and inventory management in a product-service system. Engineering Economist 59(2):91–115

Yeh R (1997) Optimal inspection and replacement policies for multi-state deteriorating systems. European Journal of Operational Research 96(2):248–259

Zhang L, Li X, Yu J (2006) A review of fault prognostics in condition based maintenance. Proceedings of SPIE , The International Society for Optical Engineering 6357 II