Exploiting Causality for Efficient Monitoring in POMDPs

Stefano V. Albrecht  
School of Informatics, The University of Edinburgh  
Edinburgh EH8 9AB, UK  
S.V.ALBRECHT@SMS.ED.AC.UK

Subramanian Ramamoorthy  
School of Informatics, The University of Edinburgh  
Edinburgh EH8 9AB, UK  
S.RAMAMOORTHY@ED.AC.UK

Abstract

POMDPs are a useful model for decision making in systems with uncertain states. One of the core tasks in a POMDP is the monitoring task, in which the belief state (i.e., the probability distribution over system states) is updated based on incomplete and noisy observations. This can be a hard problem in complex real-world systems due to the often very large state space. In this paper, we explore the idea of accelerating the monitoring task by automatically exploiting causality in the system. We consider a specific type of causal relation, called passivity, which pertains to how system variables cause changes in other variables. We present a novel monitoring method, called Passivity-based Monitoring (PM), which maintains a factored belief state representation and exploits passivity to perform selective updates over the factored beliefs. PM produces exact belief states under certain assumptions and approximate belief states otherwise, where the approximation error is bounded by the degree of uncertainty in the process. We show empirically, in synthetic processes with varying sizes and degrees of passivity, that PM is faster than two standard monitoring methods while achieving competitive accuracy. Furthermore, we demonstrate how passivity occurs naturally in a real-world system such as a multi-robot warehouse, and how PM can exploit this to accelerate the monitoring task.

1. Introduction

Partially observable Markov decision processes (POMDPs) (Thrun, Burgard, & Fox, 2005) are a useful mathematical model for decision making in systems with uncertainty about the state of the system. Rather than observing the state directly, the decision maker (or agent) observes a possibly incomplete and noisy signal which depends on the state. The signal may be incomplete in the sense that it only provides partial information about the system state, and it may be noisy in that it provides false or inaccurate information, e.g. due to unreliable sensors. Given this incompleteness and noise, it may not generally be possible to infer the system state with absolute certainty. Instead, the agent may form beliefs about the system state, based on the history of observed signals. These beliefs are represented by a belief state which is a probability distribution over the state space of the system.

One of the core tasks in a POMDP, which we here refer to as the monitoring task, is to update the belief state based on the observed signal. A sufficient condition for planning optimal actions is that the belief state be exact, in the sense that it retains all information provided by the history of signals (this is also known as sufficient statistic, (Astrom, 1965)).
In general, this can be achieved by the exact update rule, which updates each element of the belief state by propagating it through the system dynamics and conditioning it on the observed signal. However, while simple in theory, the space and time requirements of the exact update rule grow exponentially with the number of variables in the state description. Therefore, it serves mainly as a theoretical procedure which produces exact belief states, while in practice — especially for complex real-world systems with large state spaces — more efficient approximate methods are required.

Several different methods for approximate monitoring have been proposed. Gordon, Salmond, and Smith (1993) proposed a method which approximates the belief state using a set of state samples which are propagated through the system dynamics (this is now known as particle filtering, or PF). A method proposed by Boyen and Koller (1998) (henceforth referred to as BK) approximates the belief state as a product of smaller belief states defined over clusters of correlated variables. Further methods were proposed by Roy, Gordon, and Thrun (2005) who use principal component analysis to reduce the dimensionality of belief states, Poupart and Boutilier (2000) who use policy value functions to approximate belief states, and Ng, Peshkin, and Pfeffer (2002) who combined PF with BK. We postpone a detailed discussion of these and other methods to Section 2.

Each of these methods is based on the observation (e.g. (Lane & Smart, 2005)) that there is some latent structure in the process which can be exploited to render the monitoring task more tractable. For example, PF works best in systems with low degrees of uncertainty, since this means that fewer state samples are needed for acceptable approximations. On the other hand, the number of samples needed for acceptable approximations often grows drastically with the degree of uncertainty in the process (as shown in our experiments). As another example, BK works best in systems with little correlation between state variables, since this means that the variable clusters will be small and can be processed efficiently. However, if there are many variables which are strongly correlated, then BK typically becomes infeasible. Therefore, these implicit structural assumptions have to be taken into account when choosing a monitoring method for a specific system.

Another assumption which is common to the above monitoring methods is that the system dynamics are a priori known to the agent. That is, the agent knows the probabilities with which an action leads from one state to another, and the probabilities with which
observations are generated from the new state. A compact way to represent such probabilities is in the form of a dynamic Bayesian network (DBN) (Koller & Friedman, 2009), which is a probabilistic graphical model that uses parent relations to specify conditional independences between system variables (see Figure 1 for an example). These parent relations also implicitly encode causal structure (Pearl, 2000) in the system, in the sense that they specify how the system variables affect one another. To be precise, by causal structure we mean the fact that a change in one part of the system causes a change in another part of the system, either directly or indirectly through a chain of such causal effects. Our experience with such systems is that this causal structure can often be used to make the monitoring task more tractable, because it can tell us that beliefs need only be revised for certain aspects of the system state. For example, if the system variable $x_2$ in Figure 1 changes its value only in response to a change in the variable $x_1$ (i.e. $x_1$ causes the change in $x_2$), then it seems intuitive to use this causal connection when deciding whether or not to revise one’s belief about $x_2$. However, existing methods for approximate monitoring do not exploit such causal relations (we discuss this in more detail in Section 2).

It is worth pointing out that causal relations such as the one between $x_1$ and $x_2$, which we will henceforth refer to as passivity, occur naturally and frequently in many applications, especially in robotic and other physical systems (Mainzer, 2010). Intuitively, a state variable is passive if it changes its value only if (a) it is the direct target of an action, or (b) some of the variables which directly affect it change their values. The following example illustrates this in a simple robot arm:

**Example 1 (Robot arm).** Consider a robot arm with three rotational joints and a gripper, as depicted in Figure 2. The three joints are denoted by $\theta_1, \theta_2, \theta_3$ and may take any value from the discrete set $\{0^\circ, 1^\circ, \ldots, 359^\circ\}$. Assume that the value of $\theta_i$ indicates the absolute orientation of joint $i$, e.g. $\theta_i = 0^\circ$ means that joint $i$ points exactly to the right, $\theta_i = 180^\circ$ means that it points to the left, etc. Furthermore, for each joint $i$ let there be two actions, Right$_i$ and Left$_i$, which rotate the joint by $1^\circ$ clockwise and counter-clockwise, respectively.

The uncertainty in this system could be due to uncertain joint movements or unreliable sensor readings for the joint orientations.

![Figure 2: Robot arm with three rotational joints and gripper.](image)

In this example, every joint variable $\theta_i$ is passive. This is because it changes its value only if it is directly operated upon through Right$_i$ or Left$_i$ (corresponding to case (a)) or if the preceding joint variable $\theta_{i-1}$ changes (corresponding to case (b)) since a changed orientation

---

1. We mark the end of an example with a black square.
of $\theta_{i-1}$ causes a changed orientation of $\theta_i$ (recall that the orientations are absolute). Note that by extension of case (b), this also accounts for chains of causal effects, as indicated by the arrows. For example, the orientation of joint 3 changes if the orientation of joint 1 changes, since joint 1 causes joint 2 to change which in turn causes joint 3 to change.

Passivity is a persistent causal property in the sense that it is invariant to the states of the system. In other words, passive variables are always passive. However, a closer examination of case (b) in the definition of passivity reveals that passivity has a dynamic component, since whether or not one variable directly affects another variable may in general depend on the state of the system. Therefore, while passivity itself is invariant to the state, the precise causal structure may still depend on the state. The following example illustrates this:

**Example 2** (Robot arm in blocks world). Suppose we wish to use the robot arm from Example 1 to solve an instance of the blocks planning world, as shown in Figure 3. The instance consists of three blocks which are denoted by A, B, and C. The initial state of the instance is shown in Figure 3a and the goal is to stack the blocks as shown in Figure 3b. In addition to the joint variables already discussed in Example 1, a state in this system may be described by variables $X_Z$ for each block $Z \in \{A, B, C\}$ which represent their discrete horizontal positions.

In the initial state of the system, there are no variables which directly affect the variables $X_Z$, and since the blocks do not move on their own, we can conclude that they are passive.

---

2. A more complete system state would typically also include variables for the gripper and variables which describe if and how the blocks are currently stacked.
Exploiting Causality for Efficient Monitoring in POMDPs

(there are no actions which directly manipulate $X_Z$). However, consider a system state in which the gripper holds block A, as shown in Figure 3c. In this state, all three joint variables $\theta_1, \theta_2, \theta_3$ directly affect the position variable $X_A$ (indicated by the arrows), because a change in any of the joint variables causes a change in the position of A. Thus, while $X_A$ is still passive, the precise causal structure is different. Now, consider a state in which the gripper holds block B, with block A being placed on block B. In this state, the joint variables have a direct causal link to $X_B$, and in addition there is a causal link from $X_B$ to $X_A$, in that any change in $X_B$ will also cause a change in $X_A$. Thus again, while both $X_B$ and $X_A$ are still passive, the precise causal connection has changed.

The above examples were deliberately chosen to be simple in order to illustrate the idea of passivity. However, they demonstrate that passivity occurs naturally even in such simple applications. (In Section 6.2 of this paper, we show how passivity occurs in a complex example of a multi-robot warehouse system.) Furthermore, we make no claim that these examples are best modelled as POMDPs.\footnote{There are other models for observation uncertainty, including Hidden Markov Models (Stratonovich, 1960) and Kalman filters (Kalman, 1960), and they each have their own monitoring methods. However, passivity exists in all of these models because it is an inherent property of the system dynamics, not the model used to monitor the system. We focus on POMDPs because they are a general model for decision making under uncertainty and can be used for a broader variety of applications, including behaviour studies, machine vision, logistics, and medical applications (Cassandra, 1998). How could passivity be exploited to accelerate the monitoring task in the above examples?}

How could passivity be exploited to accelerate the monitoring task in the above examples? The exact update rule would update the probability for each possible system state, regardless of which action we choose. However, the fact that the system variables are passive means that some aspects of the state description may remain unchanged, depending on which action we choose. For example, if we choose to rotate joint 3, then the fact that joints 1 and 2 are passive means that they are unaffected by this action. Thus, it seems redundant to revise beliefs for the orientations of joints 1 and 2. (Example 7 in Section 5.2 shows that a saving as large as 50% can be obtained in this case.) However, this is precisely what the exact update rule does, as well as all other monitoring methods mentioned earlier.

We believe there are several features which make passivity an interesting example of a causal relation: First of all, passivity is a latent causal relation, meaning that it can be readily extracted from the system dynamics without additional annotation by an expert. (In Section 4, we give a procedure which identifies passive variables based on their dynamics.) Moreover, as seen in the above examples, passivity goes beyond the notion of conditional independence since a variable can be passive even if it depends on all other variables in the system. Furthermore, passivity is not a deterministic relation since passive variables may have any stochastic behaviour when changing their values. Finally, passivity is a relatively simple example of a causal relation, and the idea of exploiting passivity in order to accelerate the monitoring task is intuitive. Yet, to the best of our knowledge, this idea has not been formalised and explored rigorously before.

The purpose of the present paper is to formalise and evaluate the idea of automatically exploiting causal structure for efficient approximate monitoring in POMDPs, using passivity

\footnote{However, (PO)MDPs and variations thereof have been used to model blocks planning, e.g. (Geffner & Bonet, 1998; Hsiao, Kaelbling, & Lozano-Perez, 2007; Pasula, Zettlemoyer, & Kaelbling, 2007).}
as a concrete example of a causal relation. Specifically, our hypothesis is that in large systems with high degrees of passivity, this structure can be exploited to significantly accelerate the monitoring task. The idea is to maintain separate beliefs about individual aspects of the process and to exploit passivity to perform selective updates over these separate beliefs. After discussing related work in Section 2 and some technical preliminaries in Section 3, our contributions can be categorised into the following parts:

1) **Formal Definition of Passivity:** In Section 4, we give a formally concise definition of passivity and discuss various aspects of this definition. Our definition assumes that the POMDP is specified as a set of dynamic Bayesian networks (one for each action). We also discuss a non-example of passivity, by which we mean variables which appear to be passive but really are not passive. Finally, we give a simple procedure which can detect passive variables based on their dynamics.

2) **Passivity-based Monitoring:** In Section 5, we present a new method for approximate monitoring called *Passivity-based Monitoring*, or PM. Following the suggestion of Boyen and Koller (1998), PM makes the representation of the belief state more tractable by representing it as a product of smaller belief states (or belief factors) which are defined over clusters of correlated system variables. PM follows the general 2-step update procedure in which the belief state is first propagated through the system dynamics (the transition step) and then conditioned on the observed signal (the observation step). However, by keeping the belief factors separate during the update procedure, PM allows for parallel updates of the factors. The interesting novelty of PM is the way in which it performs the transition step: rather than updating all belief factors, PM updates only those factors whose variables it suspects to have changed, which is possible by exploiting passivity (to be made precise shortly). Similarly, in the observation step PM updates only those belief factors which it determines to be structurally connected with the observation, and it uses only those parts of the observation which are relevant to the belief factor, thus allowing for a more efficient incorporation of observations. We also discuss the computational complexity and error bounds of PM.

3) **Empirical Evaluation of Method:** We evaluate PM in two experimental domains: In Section 6.1, we evaluate PM in synthetic (i.e. randomly generated) processes of varying sizes and degrees of passivity. The process sizes vary from one thousand to one trillion states, and the passivity degrees vary from 25% to 100% passivity. The algorithm is compared to BK and PF, measuring accuracy in the form of relative entropy and computation time on a dual-core machine. We show that PM is faster than BK and PF while maintaining competitive accuracy. In particular, our results indicate that the computational gains grow significantly with both the degree of passivity and the size of the process. In Section 6.2, we evaluate PM in a complex simulation of a multi-robot warehouse system. The warehouse is in the style of KIVA, which is a large commercial warehouse system used by global companies such as Staples and Amazon (D’Andrea, 2012). We show how passivity occurs in this system and how PM can exploit this to accelerate the monitoring task, outperforming both BK and PF.

Finally, we discuss the strengths and weaknesses of PM in Section 7, and we conclude our work in Section 8. All proofs can be found in the appendix.
2. Related Work

Gordon et al. (1993) proposed a monitoring method,\footnote[4]{Other authors have proposed similar methods, but, to our knowledge, Gordon et al. (1993) provide the first comprehensive description.} now most commonly referred to as \textit{particle filtering} (PF), which approximates the belief state using a set of state samples (also called “particles”). The probability that the system is in state $s$ is the normalised frequency with which $s$ is contained in the state samples. PF follows a 2-step update procedure in which the current state samples are first propagated through the system dynamics and then resampled based on the probabilities with which the new state samples would have produced the observed signal. Two interesting features of PF are that it can be applied to discrete and continuous POMDPs (i.e. systems with continuous variables), and that the approximation error converges to zero as we increase the number of state samples. On the other hand, there is no theoretical guidance as to how many samples are required for good approximations, and the number of samples even for acceptable approximations can grow significantly with the degree of uncertainty in the system (as shown in our experiments in Section 6). This latter issue is one of the most severe drawbacks of PF, and several attempts have been made to address this, e.g. (Doucet, De Freitas, Murphy, & Russell, 2000).

Boyen and Koller (1998, 1999) recognised that if a system consists of several independent or weakly interacting subcomponents, then the belief state can be represented more efficiently as a product of smaller beliefs about these individual subcomponents. Their seminal contribution is to show that the approximation error due to this factored representation is essentially bounded by the degree of uncertainty (or “mixing rates”) in the system. More precisely, they prove that the relative entropy (also known as KL divergence (Kullback & Leibler, 1951)) between two belief states contracts at an exponential rate when propagated through a stochastic transition system. Furthermore, they propose a monitoring method, henceforth referred to as BK, in which the system dynamics are represented as a dynamic Bayesian network and the belief state is updated using an exact inference method such as the junction three algorithm (Lauritzen & Spiegelhalter, 1988). Since the internal “cliques” used in the junction tree algorithm may not correspond to the belief state representation of BK, a final “projection step” will typically have to be performed in which the original representation is restored. The performance of this method depends crucially on whether the relevant correlations between system variables can be captured in small clusters.

Ng et al. (2002) present a monitoring method, called \textit{factored particle filtering} (FP), which essentially combines PF and BK. FP addresses the main drawbacks of PF (many samples needed) and BK (small clusters required) by approximating the belief factors using a set of factored state samples, or “factored particles”. The samples are factored in the sense that they only assign values to the variables in the factor. This allows FP to represent belief factors which are too large for BK, and it reduces the number of samples needed due to the smaller number of variables in each factor. The authors provide different methods of updating the factored particles, but the generic idea is to first perform a “join” operation in which full particles are reconstructed from the factored particles, which are then updated as in standard PF. The updated particles are then projected down into factored particles using a “project” operation. The main drawback of FP is that these join and project operations essentially correspond to standard relational database operations, which can be very expensive. Another
way to update the factored particles is to use a junction tree, but this relies heavily on memory-efficient implementations for belief propagation.

Poupart and Boutilier (2000) argue that approximation schemes which minimise relative entropy, e.g. as suggested by Boyen and Koller (1998), are inappropriate for POMDPs because the decision quality does not necessarily depend on accurate belief states. They propose a monitoring method, called value-directed approximation, which chooses different approximation schemes for different actions so as to minimise the expected loss in decision quality (i.e. accumulated rewards). The method assumes that the POMDP has been solved exactly and that the value function is provided in the form of $\alpha$-vectors which represent the various actions in the POMDP. Based on the value function, their algorithm computes a “switching set” and “alternative plans” to determine the error bounds of approximation schemes. This is used to search for an optimal approximation scheme in a tree-based manner, where the search traverses from approximate to exact schemes. While the underlying idea of optimising decision quality is appealing, the proposed method involves a series of optimisation problems and an exhaustive tree search, which in the worst-case amounts to a quadratic increase in the solution time of the POMDP. In subsequent work, Poupart and Boutilier (2001) show how this complexity issue can be alleviated to some extent.

Roy et al. (2005) propose an integrated method in which the monitoring and control tasks are interleaved. The principal idea of their method is to reduce the dimensionality of the belief space using a dimensionality reduction method such as PCA (Jolliffe, 2005). More precisely, given a set of sample beliefs, the method tries to find a low-dimensional hyperplane which is imbedded in the belief space and such that the variance of the sampled beliefs is approximately preserved when projected onto this hyperplane. In order to reduce the reconstruction error, they propose a modified version of PCA, called “Exponential Family PCA”, which minimises the relative entropy rather than the squared error. Since the low-dimensional representation loses its convexity, conventional POMDP value iteration methods are no longer applicable. Therefore, their method uses MDP value iteration (Sutton & Barto, 1998) on a discrete belief space MDP to compute approximately optimal policies for the original POMDP. This method works well if the belief space is sparse, meaning that the probability mass is concentrated on similar or “nearby” states. In all other cases, it is difficult to find a low-dimensional hyperplane which passes through all sample beliefs. A related method based on compression was proposed by Poupart and Boutilier (2002).

We categorise monitoring methods into two groups: those which make no assumptions as to how the POMDP is solved (i.e. they are purely concerned with the monitoring task), and those methods which do make assumptions as to how the POMDP is solved. The former three methods in this section, PF, BK, FP, and our own method, PM, belong to the first group since they leave open how the POMDP is solved. On the other hand, the methods proposed by Poupart and Boutilier (2000, 2002) and Roy et al. (2005) belong to the second group since their monitoring methods assume that the POMDP is solved in some particular way. Other such methods include reachability-based methods (Washington, 1997; Hauskrecht, 2000), grid-based methods (Lovejoy, 1991; Brafman, 1997; Zhou & Hansen, 2001), and point-based methods (Pineau, Gordon, & Thrun, 2003; Smith & Simmons, 2005). The advantage of pure monitoring methods is that they can be used to monitor systems which are too complex for current solution techniques, including the multi-robot warehouse system studied in this
Moreover, they can be used independently of the solution method. On the other hand, the advantage of methods in the second group is that they can exploit synergies between the monitoring and control tasks, and explore integrated approaches such as belief space reduction, state abstraction, and value-directed methods.

None of the works discussed in this section explicitly address the question of how causal relations between system variables can be exploited to accelerate the monitoring task, or, alternatively, how the monitoring methods proposed therein implicitly benefit from causal structure. We know from our experiments that sample-based methods such as PF and FP may suffer significantly from increased degrees of passivity due to an increased sparsity in the system dynamics, and that BK indirectly benefits from the increased sparsity (though only marginally) in the junction tree algorithm. (See Sections 6.1 and 6.2 for more details.) The method proposed by Poupart and Boutilier (2000) is in some sense related to causality, because they are concerned with optimal decision quality which, ultimately, is also a question how actions are causally connected with state transitions and rewards (i.e. the well-known credit-assignment problem (Sutton & Barto, 1998)). However, their work does not address how causality can be used for efficient monitoring.

3. Preliminaries

This section provides formal definitions of POMDPs and DBNs, and establishes the basic notation used in this paper.

3.1 POMDPs, Belief States, Exact Updates

Definition 1 (POMDP). A discrete partially observable Markov decision process (POMDP), e.g. (Thrun et al., 2005), consists of

- decision maker (agent) with actions \( A = \{a_1, a_2, a_3, \ldots\} \)
- state variables \( x_1, \ldots, x_n \) with discrete domains \( X_1, \ldots, X_n \)
- observation variables \( y_1, \ldots, y_m \) with discrete domains \( Y_1, \ldots, Y_m \)
- state space \( S = X_1 \times \ldots \times X_n \)
- observation space \( O = Y_1 \times \ldots \times Y_m \)
- transition function \( T^a : S \times S \to [0, 1] \) for each \( a \in A \)
- observation function \( \Omega^a : S \times O \to [0, 1] \) for each \( a \in A \).

The process starts at time \( t = 0 \) in some initial state \( s^0 \in S \). At each time \( t \), the process is in state \( s^t \in S \) and the agent chooses an action \( a^t \in A \) based on its current belief state \( b^t \). The process then transitions into state \( s^{t+1} \in S \) with probability \( T^a(s^t, s^{t+1}) \) and the agent receives an observation \( o^{t+1} \in O \) with probability \( \Omega^a(s^{t+1}, o^{t+1}) \). For simplicity we assume that the process is time-invariant, meaning that \( T^a \) and \( \Omega^a \) are independent of \( t \). We also omit rewards since they are not relevant for our work.

5. Such systems can be controlled with domain-specific heuristics, as we do in this paper.
The belief state $b^t$ (also known as information state) represents the agent’s beliefs about what state the process might be in at time $t$ (Astrom, 1965). It takes the form of a discrete probability distribution over the state space $S$. After taking action $a^t$ and observing $o^{t+1}$, the belief state $b^t$ is updated using the exact update rule:

**Definition 2** (Exact update rule). The exact update rule is defined as

\[
\hat{b}^{t+1}(s') = \sum_{s \in S} b^t(s) T^a(s, s')
\]

(1)

\[
b^{t+1}(s') = \frac{\hat{b}^{t+1}(s') \Omega^a(s', o^{t+1})}{\sum_{s'' \in S} \hat{b}^{t+1}(s') \Omega^a(s'', o^{t+1})}.
\]

(2)

We sometimes refer to the step $b^t \rightarrow \hat{b}^{t+1}$ as the transition step and to the step $\hat{b}^{t+1} \rightarrow b^{t+1}$ as the observation step. The resulting belief state is a sufficient statistic in the sense that it summarises all information provided by the history of observed signals, and can be used to plan optimal actions (Astrom, 1965). However, the space complexity of storing exact belief states and the time complexity of updating them using the exact update rule are each exponential in the number of state variables, making it infeasible for complex real-world systems with large state spaces. This is the general motivation for our work.

**Example 3** (Robot arm as POMDP). Suppose we wish to model the robot arm from Example 1 as a POMDP. The system is described by three state variables, $\theta_1, \theta_2, \theta_3$, which take their values from the discrete domain $\Theta = \{0^\circ, ..., 359^\circ\}$. Furthermore, assume that the system has one sensor (i.e. observation) variable $\hat{\theta}_i$ for each state variable $\theta_i$, with $\hat{\theta}_i \in \Theta$. Therefore, the state and observation spaces are defined as $S = O = \Theta^3$.

The agent has access to six different actions, $A = \{\text{LEFT}_i, \text{RIGHT}_i | i = 1, 2, 3\}$. We assume that the actions have a stochastic factor, such that with probability 0.95 they produce the desired outcome while with probability 0.05 they have no effect. For instance, the transition function for $\text{LEFT}_i$ would be specified as

\[
T^{\text{LEFT}_i}(s, s') = \begin{cases} 
0.95 & s'_{\theta_j} = (s_{\theta_i} + 1^\circ) \mod 360^\circ \land \forall j \neq i : s'_{\theta_j} = s_{\theta_j} \\
0.05 & s' = s \\
0 & \text{else}
\end{cases}
\]

Similarly, the sensors are unreliable in that with probability 0.9 they report the correct values while with probability 0.1 they report random values. Thus, the observation function (for any $a \in A$) would be specified as

\[
\Omega^a(s', o) = \begin{cases} 
0.9 & \forall i : o_{\theta_i} = s'_{\theta_i} \\
0.1(|O| - 1)^{-1} & \text{else}
\end{cases}
\]

\[
\square
\]

### 3.2 Dynamic Bayesian Networks

For uniformity, we define DBNs in the context of our earlier definitions:
Definition 3 (DBN). A dynamic Bayesian network (DBN) (Koller & Friedman, 2009) for action \( a \in A \), denoted \( \Delta^a \), is an acyclic directed graph consisting of

- state variables \( X^t = \{ x^t_1, ..., x^t_n \} \) and \( X^{t+1} = \{ x^{t+1}_1, ..., x^{t+1}_n \} \)
- observation variables \( Y^{t+1} = \{ y^{t+1}_1, ..., y^{t+1}_m \} \)
- edges \( E_a \subseteq (X^t \times X^{t+1}) \cup (X^{t+1} \times X^t) \cup (X^{t+1} \times Y^{t+1}) \cup (Y^{t+1} \times X^t) \)
- probability distribution \( P_a(z | pa_a(z)) \) for each \( z \in X^t \cup Y^{t+1} \).

The variables in \( X^t \) and \( X^{t+1} \) represent the system states at times \( t \) and \( t+1 \), respectively, and the variables in \( Y^{t+1} \) represent the observation at time \( t+1 \). All variables in the DBN have the same domains as their corresponding POMDP variables, i.e. \( x^t_i \in X_i, x^{t+1}_i \in X_i \), and \( y^{t+1}_i \in Y_i \). The edges in \( E_a \) represent the inter-correlations between variables from different states (i.e. \( X^t \rightarrow X^{t+1} \)) and intra-correlations between variables from the same state (i.e. \( X^{t+1} \rightarrow X^{t+1} \)), as well as correlations between observation variables. Finally, the distributions \( P_a(z | pa_a(z)) \) specify the conditional probabilities that variable \( z \) assumes a certain value given the values of its parents \( pa_a(z) = \{ z' | (z', z) \in E_a \} \). For convenience, we also define \( pa^t_a(Z) = X^t \cap pa_a(Z) \) and \( pa^{t+1}_a(Z) = X^{t+1} \cap pa_a(Z) \), where \( pa_a(Z) = \bigcup_{z \in Z} pa_a(z) \).

The edges \( E_a \) and distributions \( P_a \) specify the dynamics of the system when executing action \( a \). Formally, the transition function \( T^a \) and observation function \( \Omega^a \) for action \( a \) are defined as

\[
T^a(s, s') = \prod_{i=1}^n P_a(x_i^{t+1} = s'_i | pa_a(x_i^{t+1}) \leftarrow (s, s'))
\]

\[
\Omega^a(s', o) = \prod_{j=1}^m P_a(y_j^{t+1} = o_j | pa_a(y_j^{t+1}) \leftarrow (s', o))
\]

where we use the notation \( pa_a(x_i^{t+1}) \leftarrow (s', s'^{t+1}) \) to denote the tuple in which the parents of \( x_i^{t+1} \) in \( X^t \) and \( X^{t+1} \) respectively, assume values from \( s \) and \( s' \). Formally, if \( x_i^t \in pa^t_a(x_i^{t+1}) \) and \( x_i^{t+1} \in pa^{t+1}_a(x_i^{t+1}) \), then \( x_i^t = s_i \) and \( x_i^{t+1} = s'_i \). Similarly, we use the abbreviated notation \( pa_a(y_j^{t+1}) \leftarrow (s', o) \) to denote the tuple in which the parents of \( y_j^{t+1} \) in \( X^{t+1} \) and \( Y^{t+1} \), respectively, assume values from \( s' \) and \( o \).

Example 4 (DBN representation of robot POMDP). We can represent the POMDP from Example 3 as a set of DBNs, where we have one DBN \( \Delta^a \) for each action \( a \in A \). The state and observation variables in the DBNs are \( X^t = \{ \theta_1^t, \theta_2^t, \theta_3^t \} \), \( X^{t+1} = \{ \theta_1^{t+1}, \theta_2^{t+1}, \theta_3^{t+1} \} \), and \( Y^{t+1} = \{ \hat{\theta}_1^{t+1}, \hat{\theta}_2^{t+1}, \hat{\theta}_3^{t+1} \} \). To make our example more realistic, let us assume that the joint orientations are bounded relative to the orientation of the immediately preceding joint (e.g. in the form of a cone), where the first joint is bounded relative to the ground. This means that the joint movement depends on its own as well as the preceding joint orientation, as shown in Figure 4. Moreover, the joint orientations are correlated (i.e. edges within \( X^{t+1} \)) such that no joint can exceed the bound given by the preceding joint. Finally, the observation variables depend solely on the corresponding joint variable.\(^6\)

\(^6\) There are typically not just one but several ways to specify the edges in a DBN, and it is the task of the system designer to find a specification which is both natural and practical. We often find it useful to first
The actions in this example would differ in their variable distributions $P_a$. For analytical purposes, it is often convenient to represent the distributions in tabular form. This is referred to as conditional probability table, or CPT (Russell & Norvig, 2002). However, for complex systems the variable distributions may have to be represented using a more compact method. In this work, we make no assumptions as to the concrete choice of representation, though we do assume that the distributions are represented correctly (that is, there are no inaccuracies due to approximate representations).

3.3 Notational Conventions and Additional Definitions

In this work, we use DBNs to specify the dynamics of a POMDP. Specifically, a DBN $\Delta^a$ specifies how the system variables $x_i$ and $y_i$ affect each other when executing action $a$. Given the close relation of these models, it will be useful to establish some notational conventions:

- Whenever we refer to a system variable $x_i$ or $y_j$, this is to be read in the general context of the modelled system (i.e. the POMDP), whereas whenever we refer to the variables $x_i^t, x_i^{t+1}$, or $y_j^{t+1}$, this is to be read in the specific context of a DBN $\Delta^a$.  

- The variables $x_i, x_i^t, x_i^{t+1}$ generally all refer to the same state variable and have the same discrete domain $X_i$. Given a state $s \in S$, we use the notation $s_i \in X_i$ to refer to the value of $x_i$ (or $x_i^t, x_i^{t+1}$) in state $s$.

- The variables $y_j$ and $y_j^{t+1}$ generally refer to the same observation variable and have the same discrete domain $Y_j$. Given an observation $o \in O$, we use the notation $o_j \in Y_j$ to refer to the value of $y_j$ (or $y_j^{t+1}$) in observation $o$.

Furthermore, it will be convenient to define the following:

- specify the correlations between variables (i.e. edges in $X^{t+1}$) such that they can be used to “test” for invalid states. Then, further edges from $X^t$ can be added as required.
The binary order $<$ is defined over $X^t \cup X^{t+1}$ such that $x_i^t < x_j^t$ and $x_i^{t+1} < x_j^{t+1}$ for all $1 \leq i < j \leq n$, and $x_i^t < x_j^{t+1}$ for all $1 \leq i, j \leq n$.

Given a set $Z \subseteq X^t \cup X^{t+1}$, we write $Z^<$ to denote the ordered tuple $(z_{i_1}, ..., z_{i_{|Z|}})$ in which each $x_i^t \in Z$ or $x_i^{t+1} \in Z$ is represented by one $z_{i_i}$ with $i_i = i$, and such that $z_{i_i} < z_{i_{i+1}}$ (that is, $Z^<$ contains all variables of $Z$, ordered by $<$).

Given the ordered tuple $Z^< = (z_{i_1}, ..., z_{i_{|Z|}})$, let $X_{z_i}$ be the discrete domain of variable $z_{i_i}$. Then, we define the set $S(Z) = X_{z_{i_1}} \times \cdots \times X_{z_{i_{|Z|}}}$ which contains all value tuples for the variables in $Z$.

Given a value tuple $s_Z = (s_{i_1}, ..., s_{i_{|Z|}}) \in S(Z)$, we use the notation $Z \leftrightarrow s_Z$ as an abbreviation for $z_{i_i} = s_{i_i}$ for each $z_{i_i} \in Z$ (i.e. the variables in $Z$ assume their corresponding values from $s_Z$).

4. Passivity

This section introduces a formal definition of passivity, which will then be used as the basis for the remainder of this paper.

4.1 Formal Definition

Recall from Section 1 that a state variable $x_i$ is called passive if it changes its value only if (a) it is the direct target of an action (i.e. the action directly manipulates the value of $x_i$), or (b) some of the variables which directly affect $x_i$ change their values. Case (a) is implicit in the specification of the actions, in the sense that if action $a$ directly manipulates $x_i$, then the variable distribution $P_a$ of $x_i$ will reflect the intended manipulation of $x_i$. For instance, in Example 1 the action RIGHT3 directly manipulates the variable $\theta_3$ such that its value rotates clock-wise by $1^\circ$. Therefore, given that the system dynamics are specified as a set of DBNs $\Delta^a$, one for each action $a$, we can reduce the definition of passivity to case (b) such that the passivity of a variable is defined with respect to the actions in the system.

Formally, we define passivity as follows:

**Definition 4** (Passivity). Let action $a$ be given by a DBN $\Delta^a$. A state variable $x_i$ is called passive in $\Delta^a$ if there exists a set $\Phi_{a,i} \subseteq \text{pa}_a(x_i^{t+1}) \setminus \{x_i^t\}$ such that:

(i) $\forall x_j^t \in \Phi_{a,i} : \left( x_j^{t+1} \right) \in E_a$

and

(ii) for any two states $s^t$ and $s^{t+1}$ with $T_a(s^t, s^{t+1}) > 0$:

$$\left( \forall x_j^t \in \Phi_{a,i} : s_j^t = s_j^{t+1} \right) \Rightarrow s_i^t = s_i^{t+1}$$

(5)

A state variable which is not passive is called active.

The set $\Phi_{a,i}$ corresponds to the set of variables described in case (b) above: it contains all those variables which directly affect $x_i$ (i.e. they are parents of $x_i^{t+1}$ in $X^t$) such that $x_i$ may change its value only if any of the variables in $\Phi_{a,i}$ change their values. We will sometimes
say that a variable $x_i$ is passive in $\Delta^a$ with respect to another variable $x_j$ if it is the case that $x_j \in \Phi_{a,i}$. Furthermore, we will omit "in $\Delta^a$" if it is obvious from the context.

Clause (i) in Definition 4 requires that $x_i$ is intra-correlated with the variables in $\Phi_{a,i}$, specifically that there is an edge from $x_j$ to $x_i$ for all $x_j \in \Phi_{a,i}$. As an example, see Figure 1 in which we assumed that the variable $x_2$ was passive with respect to the variable $x_1$. (We will discuss the purpose of this clause in the next subsection.) Clause (ii) defines the core semantics of passivity by requiring that $x_i$ remains unchanged if all variables in $\Phi_{a,i}$ remain unchanged. Note that this means that the distribution $P_a$ for $x_i$ may specify any deterministic or stochastic behaviour if the variables in $\Phi_{a,i}$ change their values. This includes that $x_i$ may not change its value at all.

A state variable can be passive even if it has no parents in $X^t$, or no parents other than itself. In this case, the set $\Phi_{a,i}$ would be empty and the clause (i) as well as the premise in (5) would trivially hold true. However, such a variable can only be passive if it does not change its value under any circumstances. In other words, it would have to be a constant, in which case one should consider removing the variable from the state description in order to reduce computational costs.

According to Definition 4, any variable which is directly manipulated by an action would be called an active variable. This is since these variables do not depend on a change in other variables to "trigger" their own change (instead, the action triggers the change directly). However, there may be active variables other than the ones directly targeted by the action. Such variables can be considered as changing on their own. For example, if we monitor a ball rolling down a slope, then the velocity of the ball would be an active variable since it changes its value without any triggering by an action or changes in other variables.

### 4.2 Non-Example of Passivity

What is the purpose of clause (i) in the definition of passivity? After all, and as discussed previously, clause (ii) captures the core idea of passivity which is that a variable may only change its value if any of the variables with respect to which it is passive change their values. However, while it may seem intuitive that clause (ii) be sufficient for passivity, there are in fact systems in which clause (ii) alone does not suffice. In other words, clause (ii) is necessary but not sufficient for passivity. We illustrate this in the following example:

**Example 5** (Non-example of passivity). Consider a system with two binary state variables, $x_1, x_2$, and a single action, $a$, shown in Figure 5. (For the purposes of this example we omit the observation variables.) The dynamics of the system are such that $x_1^{t+1}$ takes the value of $x_2^t$ and $x_2^{t+1}$ takes the value of $x_1^t$ (i.e. $x_1$ and $x_2$ swap their values at each time step).

In this system, both state variables satisfy clause (ii) of Definition 4: If we set $x_0^0 = x_2^0$ (i.e. same initial values), then $T^a(s^t, s^{t+1})$ is positive only for states $s^t = s^{t+1}$, and hence (5) is true. If we set $x_0^0 \neq x_2^0$, then $T^a(s^t, s^{t+1})$ is positive only for states $s^t, s^{t+1}$ with $s_i^t \neq s_i^{t+1}, i \in \{1, 2\}$, and hence (5) is trivially true since its premise is false.

Despite satisfying clause (ii), the state variables $x_1$ and $x_2$ from Example 5 are in fact not passive, for the following two reasons: Firstly, passivity is a causal relation and as such it must imply a causal order (Pearl, 2000). However, there is no causal order between $x_1$ and $x_2$, because there there is no edge between $x_1^{t+1}$ and $x_2^{t+1}$. Secondly, passivity means that a
variable changes only if another variable with respect to which it is passive (i.e. a variable in $\Phi_{a,i}$) changes. In other words, whether or not a passive variable $x_i$ changes depends on both the past values of $\Phi_{a,i}$ (at time $t$) and the new values of $\Phi_{a,i}$ (at time $t+1$). However, the variables in Example 5 only depend on the values at time $t$, thus their own values at time $t+1$ are predetermined and do not depend on whether the variables in $\Phi_{a,i}$ change.

The first issue, namely that of the causal order, can be addressed by adding the corresponding edges in $X^{t+1}$. For instance, in Example 5 we could add an edge from $x_{t+1}^1$ to $x_{t+1}^2$ to establish a causal order. However, this does not generally solve the second issue, which is that every passive variable $x_i$ must depend both past and new values of the variables in $\Phi_{a,i}$.

In other words, $x_i$ must be both inter-correlated as well as intra-correlated with the variables in $\Phi_{a,i}$. The former is given by definition (since every variable in $\Phi_{a,i}$ is a parent of $x_i$) and the latter is precisely what is required by clause (i) in Definition 4. Therefore, clauses (i) and (ii) together define formally what we mean by passivity.

4.3 Detecting Passive Variables

As mentioned in Section 1, passivity is a latent causal property in the sense that it can be extracted from the system dynamics without additional information, and with no additional assumptions regarding the representation of variable distributions. In order to determine if a variable $x_i$ is passive in $\Delta^a$, one has to find a set $\Phi_{a,i}$ such that both clauses of Definition 4 are satisfied. A simple procedure which does this for any representation of the variable distributions is given in Algorithm 1. The algorithm takes as inputs a variable $x_i$ and a DBN $\Delta^a$, and checks whether $x_i$ is passive in $\Delta^a$ by searching for a set $\Phi_{a,i}$ which satisfies both clauses of Definition 4. Note that the power set $\mathcal{P}$ in line 3 includes the empty set $\emptyset$, hence it also accounts for $\Phi_{a,i} = \emptyset$. Lines 7 to 11 check if clause (i) is satisfied while lines 12 to 18 check if clause (ii) is satisfied. Line 15 essentially checks if (5) holds true. If both clauses are satisfied, then $x_i$ is passive in $\Delta^a$ with respect to the variables in $\Phi_{a,i}$, and the algorithm returns the set $\Phi_{a,i}$. Otherwise, the algorithm returns a logical false.7

Note that the set $\Phi_{a,i}$ is not necessarily unique. For example, consider a variable $x_1$ which is passive in $\Delta^a$ with respect to variables $x_2$ and $x_3$, i.e. $\Phi_{a,1} = \{x_2, x_3\}$, and assume that $x_2$ changes if and only if $x_3$ changes (i.e. they change at the same time). Then, it is easy to verify that $\Phi'_{a,1} = \{x_2\}$ and $\Phi''_{a,1} = \{x_3\}$ also satisfy clauses (i) and (ii), and hence $\Phi_{a,1}, \Phi'_{a,1}, \Phi''_{a,1}$ are all valid sets under our definition of passivity. The guiding principle in such cases is

---

7. Strictly speaking, Algorithm 1 checks for a property which is stronger than passivity because it does not check for $T^a(s^t, s^{t+1}) > 0$ (cf. clause (ii)) in line 14. However, the algorithm can be modified to include such a check. We omit this in our exposition in order to highlight the core ideas behind the algorithm.
Algorithm 1 Passive($x_i, \Delta^a$)

1: **Input:** variable $x_i$, DBN $\Delta^a$
2: **Output:** $\Phi_{a,i}$ if $x_i$ is passive in $\Delta^a$, else **false**
3: $Q \leftarrow \text{ORDEREDQUEUE}(P(pa_a^t(x_i^{t+1})))$ // in ascending order of $|\Phi_{a,i}|$
4: while $Q \neq \emptyset$ do
5: $\Phi_{a,i} \leftarrow \text{NEXTELEMENT}(Q)$
6: $Q \leftarrow Q \setminus \{\Phi_{a,i}\}$
7: for all $x_j^t \in \Phi_{a,i}$ do
8: if $(x_j^{t+1}, x_i^{t+1}) \notin E_a$ then
9: Go to line 4
10: end if
11: end for
12: $\Psi_{a,i} \leftarrow pa_a(x_i^{t+1}) \setminus \Phi_{a,i}$
13: $\Phi_{a,i}^{t+1} \leftarrow \{x_j^{t+1} | x_j^t \in \Phi_{a,i}\}$
14: for all $s_\Phi \in S(\Phi_{a,i}),$ $s_\Psi \in S(\Psi_{a,i}),$ $s_i \in X_i$ do
15: if $P_a(x_i^{t+1} = s_i | x_i^t = s_i, \Phi_{a,i} \leftarrow s_\Phi, \Phi_{a,i}^{t+1} \leftarrow s_\Phi, \Psi_{a,i} \leftarrow s_\Psi) < 1$ then
16: Go to line 4
17: end if
18: end for
19: return $\Phi_{a,i}$
20: end while
21: return **false**

*Occam’s razor*, which, intuitively speaking, states that the simplest explanation suffices. In our case, this means that it suffices to know the smallest set $\Phi_{a,i}$ in terms of the cardinality $|\Phi_{a,i}|$. (Hence, line 3 in Algorithm 1 sorts the queue $Q$ in ascending order of $|\Phi_{a,i}|$.) The rationale is that if there exist multiple causal explanations for a passive variable $x_i$, then the one involving the fewest key variables is to be favoured since it reduces (compared to the alternative explanations) the number of cases in which we would have to revise our beliefs about $x_i$. In our earlier example, if we accept $\Phi_{a,1}$ as a causal explanation for $x_1$, then we would have to revise our beliefs for $x_1$ every time $x_2$ or $x_3$ may have changed their values. However, if we accept $\Phi'_{a,1}$ as a causal explanation, then we would have to revise our belief for $x_1$ only if $x_2$ may have changed its value. This difference will become more obvious in Section 5.2, which explains how passivity can be exploited to reduce computational costs.

The time complexity of Algorithm 1 is exponential in the worst case, in which $x_i$ is not passive. Specifically, the time requirements of line 4 grow exponentially with the number
of parents of $x_i$ in $X^t$, and the time requirements of line 14 grow exponentially with the cardinality of $\Phi_{a,i}$ and $\Psi_{a,i}$. However, these time requirements can be reduced significantly when committing to specific representations for the variable distributions $P_a$. For example, if the distributions are represented in tabular form, then one can utilise arrays of indices to perform sweeping tests of (5), i.e. line 15. Moreover, it is important to realise that the algorithm needs to be performed only once for each variable, prior to the start of the process. This is since passivity (as discussed in Section 1) is invariant of the system states. In other words, if a variable is passive in $\Delta^a$, then it will always be passive in $\Delta^a$. Therefore, it suffices to check once in advance for passivity.

5. Passivity-based Monitoring

This section presents our monitoring algorithm, called Passivity-based Monitoring, or PM for short. As discussed in Section 3, we assume that the process is specified as a set of dynamic Bayesian networks which contains one DBN $\Delta^a$ for each action $a \in A$. Therefore, whenever we refer to an action $a$ (e.g. $T^a$, $\Omega^a$, $P_a$, $p_a$, etc.), this is assumed to be in the context of $\Delta^a$.

PM follows the general two-step update procedure in which the belief state is first propagated through the system dynamics (transition step) and then conditioned on the observed signal (observation step). Thus, it is natural to divide the exposition of PM into three parts: (1) the belief state representation; (2) the transition step; and (3) the observation step. These are discussed in Sections 5.1, 5.2, and 5.3, respectively. A summary of PM is given in Section 5.4. We also discuss the computational complexity of PM in Section 5.5, and the error bounds of PM in Section 5.6.

5.1 Belief State Representation

Recall from Section 1 that the principal idea behind PM is to maintain separate beliefs about individual aspects of the process, and to exploit passivity in order to perform selective updates over these separate beliefs. The union of all individual aspects constitutes a complete state description of the process. Therefore, the belief state can be represented as the product of all separate beliefs about the individual aspects.

We capture the informal notion of “individual aspects” formally in the form of clusters, which are defined as follows:

**Definition 5 (Cluster).** A clustering of $X^{t+1}$ is a set $C = \{C_1, ..., C_K\}$ which satisfies $\forall k : C_k \subseteq X^{t+1}$ and $C_1 \cup ... \cup C_K = X^{t+1}$. We refer to the elements $C_k$ of $C$ as clusters.

The underlying idea behind the concept of clusters is that the variables in a cluster $C_k$ are connected in some important sense. Specifically, if two or more variables are in a common cluster, then there exists some relation between these variables regarding the likelihood of values which they may assume. In other words, the variables are correlated in $X^{t+1}$.

The number $K$ and the concrete choice of clusters $C_k$ can be specified by the user or generated automatically. For example, they may be specified by a domain exert which is familiar with the structure of the modelled system, or generated automatically using methods such as the ones described in Section 6.1. It should be stressed, however, that in order to reduce computational costs it is advisable to follow the general rule “as small as possible, as large as necessary” when choosing clusters (see Section 5.5 for a discussion
about computational complexity). Therefore, if two variables are strongly correlated, then they should presumably be in a common cluster, whereas if they are not or only weakly correlated (“weakly” meaning that the correlation can be ignored safely), then they should be in separate clusters in order to reduce computational costs. This is illustrated in the following example:

**Example 6 (Clusters in robot DBN).** Recall the robot arm DBN from Example 4, specifically Figure 4. One way to cluster the state variables in $X_{t+1}$ is given by the three clusters $C_1 = \{\theta_{t+1}^1\}$, $C_2 = \{\theta_{t+1}^2\}$, $C_3 = \{\theta_{t+1}^3\}$, as shown in Figure 6a. This clustering is most efficient since it minimises the size of each cluster. However, the clusters fail to capture the important correlation that the joint orientation $\theta_i$ is restricted by the preceding joint orientation $\theta_{i-1}$. Another way to cluster the state variables is given by the single cluster $C_1 = \{\theta_{t+1}^1, \theta_{t+1}^2, \theta_{t+1}^3\}$, as shown in Figure 6b. This clustering captures all correlations between variables. However, this is the largest possible cluster and, therefore, the least efficient one. A compromise is given by the two clusters $C_1 = \{\theta_{t+1}^1, \theta_{t+1}^2\}$, $C_2 = \{\theta_{t+1}^2, \theta_{t+1}^3\}$, which are shown in Figure 6c. This clustering captures the correlation of the joint orientations with the immediately preceding joint orientations, and it is more efficient than the previous clustering since it has smaller clusters.

Given the definition of clusters, we capture the informal notion of “separate beliefs” in the form of belief factors:

**Definition 6 (Belief factor).** Given a cluster $C_k$, the corresponding belief factor $b_k$ is a discrete probability distribution over the set $S(C_k)$.

Intuitively, a belief factor $b_k$ represents the agent’s beliefs as to the likelihood of values for the variables in the corresponding cluster $C_k$. An analogy to this is to view a belief factor as a “smaller” belief state, and to view $b$ as the “full” belief state which is a combination of the smaller belief states. However, to distinguish the two, we refer to $b$ simply as the belief state and to $b_k$ as a belief factor.
Finally, given the clusters $C_k$ and their corresponding belief factors $b_k$, the belief state $b$ is represented in factored form as

$$b(s) = \prod_{k=1}^{K} b_k(s_k)$$

where we use the notation $s_k$ to refer to the tuple $(s_i)_{x_i \in C_k}$. (E.g., if $C_k = \{x_2^{t+1}, x_3^{t+1}\}$ and $s = (s_1, s_2, s_3, s_4)$, then $s_k = (s_2, s_3)$.)

### 5.2 Exploiting Passivity in the Transition Step

In order to perform selective updates over the belief factors $b_k$, we require a procedure which performs the transition step independently for each factor. We obtain such a procedure by introducing two assumptions which allow us to modify the transition step (1) of the exact update rule. The assumptions guarantee that the transition step is performed exactly, in the sense of (1). However, as we will discuss shortly, the assumptions can be violated to obtain approximate belief states.

The first assumption, (A1), states that the clusters must be uncorrelated (i.e. there are no edges in $X^{t+1}$ between clusters), and the second assumption, (A2), states that the clusters must be disjoint. Formally, these are defined as follows:

(A1) \( \forall a : x_i^{t+1} \in C_k \rightarrow pa_a^{t+1}(x_i^{t+1}) \subseteq C_k \)

(A2) \( \forall k \neq k' : C_k \cap C_k' = \emptyset \)

Note that neither assumption implies the other. That is, it may be the case that (A1) is satisfied while (A2) is violated, and vice versa. Given that (A1) and (A2) both hold, we can reformulate (1) to

$$\hat{b}_k^{t+1}(s_k') = \eta_1 \sum_{\bar{s} \in \bar{S}(pa_a^{t+1}(C_k))} T^a_k(\bar{s}, s_k') \prod_{k' : C_k \cap pa_a^{t+1}(C_k) \neq \emptyset} b_{k'}^{t}(\bar{s}_{k'})$$

where $\eta_1$ is a normalisation constant and

$$T^a_k(\bar{s}, s_k') = \prod_{x_i^{t+1} \in C_k} P_a(x_i^{t+1} = (s_k')_i \mid pa_a(x_i^{t+1}) \leftrightarrow (\bar{s}, s_k')) .$$

This procedure performs the transition step independently for each belief factor $b_k$, hence they can be updated in any order and in parallel.

Assumption (A1) is what allows us to bring (1) into a form which updates the belief factors $b_k$ independently of each other. Specifically, (A1) allows us to define the cluster-based transition function $T^a_k$, which in turn enables the summation in (1). Assumption (A2), on the other hand, guarantees that the product in (1) is correct. In particular, it may be the case that $|\bar{s}_{k'}| < |C_{k'}|$ (i.e. there are fewer elements in $\bar{s}_{k'}$ than in $C_{k'}$) if there are variables in $C_{k'}$ which are not in $pa_a^{t+1}(C_k)$ (i.e. $x_i^{t+1} \in C_{k'}$ but $x_i^{t+1} \notin pa_a^{t+1}(C_k)$). In such cases, $b_{k'}^{t}$ is taken

---

8. This also has the advantage that the belief factors can be updated in parallel, which is a useful feature considering that many platforms use parallel processing techniques.
to be the marginal distribution over the variables in $C_k' \cap pa'_{a,t}(C_k)$, where (A2) guarantees that the marginalisation introduces no errors.

As mentioned previously, each assumption may be violated to obtain approximate belief states. However, there is an important distinction between (A1) and (A2) in this regard: If (A2) is violated, then (1) is still well-defined (in the sense that it can still be executed), except that the product in (1) may degrade the accuracy of the results. This is in contrast to (A1), which is a structural requirement of $T^a_k$ in the sense that $T^a_k$ is ill-defined without (A1). This is since, if (A1) is violated, then the variables in $C_k$ may have parents in $X^{t+1}$ which are not in $C_k$, in which case $pa_a(x^{t+1}_i) \leftrightarrow (s, s'_k)$ would be ill-defined. Thus, if (A1) is violated, we have to enforce it by modifying the distributions $P_a$ of all $x^{t+1}_i \in C_k$ to marginalise out all variables in $pa'_{a,t}(x^{t+1}_i)$ which are not in $C_k$, for all clusters $C_k$. This means that each variable has a separate distribution for every cluster which contains the variable, thus possibly introducing an approximation error.

Given the modified transition step (1), we can exploit passivity to decide which of the belief factors $b_k$ should be updated. Recall from Section 4.1 that a variable $x_i$ is passive in $\Delta^a$ if there is a set $\Phi_{a,i}$ of variables such that $x_i$ changes its value only if any of the variables in $\Phi_{a,i}$ change their values. This causal connection can be used to decide whether or not the variables of the values in a cluster $C_k$ may have changed, in which case the corresponding belief factor $b_k$ should be updated. Theorem 1 provides the formal foundation:

**Theorem 1.** If (A1) and (A2) hold, and if all $x^{t+1}_i \in C_k$ are passive in $\Delta^a$, then

$$\forall s \in S : \hat{b}^{t+1}_k(s_k) = b^{t+1}_k(s_k).$$

**Proof.** Proof in Appendix A. \qed

Theorem 1 states that if the clusters $C_1, ..., C_K$ are disjoint and uncorrelated, and if all variables in cluster $C_k$ are passive in $\Delta^a$, then the transition step for the corresponding belief factor $b^t \rightarrow \hat{b}^{t+1}$ can be omitted without loss of information.

How does Theorem 1 translate into situations in which (A1) or (A2), or both, are violated? The key assumption is again (A1), which states that the clusters must be uncorrelated. As discussed earlier, we can enforce this by modifying the variable distributions $P_a$ in each cluster. However, if a passive variable $x^{t+1}_i \in C_k$ is correlated with a (passive or active) variable $x^{t+1}_j \in C_{k'}$, where $x^{t+1}_j \in pa_a^{t+1}(x^{t+1}_i)$, then marginalising out $x^{t+1}_j$ in the distribution $P_a$ of $x^{t+1}_i$ will typically cause $x^{t+1}_i$ to lose its passivity, in the sense that it would no longer satisfy the clauses in Definition 4. Consequently, we would always have to perform the transition step for $C_k$, even if the unmodified variables in $C_k$ are all passive. This is problematic not only because of the unnecessary computations, but also because the modified distributions will introduce an error every time the transition step is performed.

To alleviate this effect, one can check if there is a chance that the *unmodified* variables in the cluster would change their values. It can be shown that this is the case whenever there is a *causal path* from any active variable to a variable in the cluster:

**Definition 7** (Causal path). A *causal path* in $\Delta^a$, from an active variable $x^{t+1}_i$ to another variable $x^{t+1}_j$, is a sequence $(x^{(1)}, x^{(2)}, ..., x^{(Q)})$ such that $x^{(1)} = x^{t+1}_i, x^{(q)} = x^{t+1}_j$, and for all $q < Q$:

(i) $x^{(q)} \in X^{t+1}$
(ii) \((x^{(q)}, x^{(q+1)}) \in E_a\)
(iii) \(x^{(q+1)}\) is passive in \(\Delta^a\) with respect to \(x^{(q)}\)

Thus, as a general update rule, we can omit the transition step \(b^t_k \rightarrow \hat{b}^{t+1}_k\) if all unmodified variables in cluster \(C_k\) are passive in \(\Delta^a\), and if there is no causal path from any active variable in \(\Delta^a\) to any variable in \(C_k\). We demonstrate this in the following example:

**Example 7** (Update rule in robot arm DBN). Let us again consider the robot arm from the previous examples. Figure 7 shows a DBN which implements the action \(\text{Right}_3\). This action rotates joint 3 of the robot arm by \(1^\circ\) clock-wise (i.e. the joint orientation \(\theta^{t+1}_3\) is a direct target of the action). Therefore, the variable \(\theta^{t+1}_3\) is active (shown as dashed circle) while the variables \(\theta^{t+1}_1\) and \(\theta^{t+1}_2\) are passive.

![Figure 7: Robot arm DBN implementing the action \(\text{Right}_3\). Solid white circles mark passive variables, dashed white circles mark active variables, and solid grey circles mark observation variables. The coloured ellipses represent the clusters \(C_1\) and \(C_2\).](image)

We use the clustering \(C_1 = \{\theta^{t+1}_1, \theta^{t+1}_2\}\), \(C_2 = \{\theta^{t+1}_2, \theta^{t+1}_3\}\) for reasons given in Example 6. Since \(\theta^{t+1}_1\) is a parent of \(\theta^{t+1}_2\), PM will have to enforce assumption (A1) by marginalising \(\theta^{t+1}_1\) out of the variable distribution \(P_a\) of \(\theta^{t+1}_2\) in cluster \(C_2\). While the modified variable distribution loses the passivity property (both clauses of Definition 4 are violated), the unmodified distribution of \(\theta^{t+1}_1\) is still passive.

When performing the transition step, PM has to update the belief factor \(b_2\) because the corresponding cluster \(C_2\) contains the active variable \(\theta^{t+1}_3\). However, since all variables in cluster \(C_1\) are passive (there are no modified variables in \(C_1\)), and since there is no causal path from \(\theta^{t+1}_3\) to any variable in \(C_1\), PM can omit the update for the belief factor \(b_1\). Intuitively, this makes sense since a change in the orientation of joint 3 cannot cause a change in the orientations of the preceding joints. Note that this corresponds to a saving of 50% in the transition step.

Algorithm 2 defines a procedure which utilises this rule to find clusters for which the transition step can be skipped. The algorithm takes as inputs a set \(C\) of clusters and a
Algorithm 2 SkippableClusters(\(\mathcal{C}, \Delta^a\))

1: **Input:** set of clusters \(\mathcal{C} = \{C_1, \ldots, C_K\}\), DBN \(\Delta^a\)

2: **Output:** set of clusters \(\mathcal{C}^* \subseteq \mathcal{C}\) which can be skipped in transition step

3: \(\mathcal{C}^* \leftarrow \mathcal{C}\)

4: \(Q \leftarrow \text{OrderedQueue}(X^{t+1})\)

5: while \(\mathcal{C}^* \neq \emptyset \land Q \neq \emptyset\) do

6: \(x_i^{t+1} \leftarrow \text{NextElement}(Q)\)

7: \(Q \leftarrow Q \setminus \{x_i^{t+1}\}\)

8: if \(\neg \text{Passive}(x_i, \Delta^a)\) then

9: \(\mathcal{C}^* \leftarrow \mathcal{C}^* \setminus \{C_k \in \mathcal{C}^* \mid x_i^{t+1} \in C_k\}\)

10: for all \(x_j^{t+1} \in Q\) do

11: if \(\text{CausalPath}(x_i^{t+1}, x_j^{t+1}, \Delta^a)\) then

12: \(\mathcal{C}^* \leftarrow \mathcal{C}^* \setminus \{C_k \in \mathcal{C}^* \mid x_j^{t+1} \in C_k\}\)

13: \(Q \leftarrow Q \setminus \{x_j^{t+1}\}\)

14: end if

15: end for

16: end if

17: end while

18: return \(\mathcal{C}^*\)

DBN \(\Delta^a\), and returns a set \(\mathcal{C}^*\) of skippable clusters. It essentially searches through all active variables \(x_i^{t+1}\) in \(\Delta^a\) and removes all clusters \(C_k\) from \(\mathcal{C}\) which contain variables to which there is a causal path from \(x_i^{t+1}\). The function \(\text{OrderedQueue}(X^{t+1})\) returns an ordered queue \(Q\) with all variables in \(X^{t+1}\). The performance of Algorithm 2 depends on the order of the queue. In our experiments, we obtained good performance by ordering the variables in descending order of their number of outgoing edges. The function \(\text{NextElement}(Q)\) returns the next element in the queue; the function \(\text{Passive}(x_i, \Delta^a)\) was defined in Algorithm 1; and the function \(\text{CausalPath}(x_i^{t+1}, x_j^{t+1}, \Delta^a)\) returns a logical true if and only if there is a causal path from \(x_i^{t+1}\) to \(x_j^{t+1}\) in \(\Delta^a\). Note that, given the invariance of passivity to system states (cf. Section 4.1), it suffices to call Algorithm 2 only once (in advance or as needed) to determine which of the clusters to omit in the transition step.

---

9. A simple way to implement this function is to modify a standard graph search method (such as breath-first search) to check for (iii) in Definition 7, and to apply it to the variables in \(X^{t+1}\) with edges \(E_a\) from \(\Delta^a\).
5.3 Efficient Incorporation of Observations

PM can perform the observation step similarly to the exact update rule (2), which conditions the propagated belief state \( \hat{b}_{k'}^{t+1} \) on the observed signal \( \sigma_t^{t+1} \) to obtain a fully updated belief state \( b_t^{t+1} \). However, given the factored belief state representation used by PM, we require a procedure which respects this factorisation in the observation step. Given that the assumptions (A1) and (A2) both hold, we can bring (2) into a form which updates the belief factors \( b_k \) independently of each other

\[
\begin{align*}
\hat{b}_{k}^{t+1}(s_k) &= \eta_2 \sum_{s' \in S(pa_{a_t}^{t+1}(Y^{t+1}))} \Omega_a(s,s') \prod_{k' \neq k: C_{k'} \cap pa_{a_t}^{t+1}(Y^{t+1}) \neq \emptyset} \hat{b}_{k'}^{t+1}(s_{k'})
\end{align*}
\]

where \( \eta_2 \) is a normalisation constant. Note that, analogously to (6), if there are variables in \( C_{k'} \) which are not in \( pa_{a_t}^{t+1}(Y^{t+1}) \), then \( \hat{b}_{k'}^{t+1} \) is taken to be the marginal distribution over \( C_{k'} \). Assumption (A2) guarantees that the marginalisation introduces no errors. If (A1) and (A2) both hold, then the transition step (6) and observation step (7) produce exact belief states in the sense of (1) and (2), regardless of how many clusters were skipped in the transition step (cf. Theorem 1).

The observation step (7) updates all belief states and uses all observation variables in the process. In other words, it ignores the internal structure of the observation variables. However, it is clear that if the variables in a cluster \( C_k \) are marginally independent of the observation variables \( Y^{t+1} \) (this can be determined using d-separation (Geiger, Verma, & Pearl, 1990), or simply by checking if there is a directed path from \( Y^{t+1} \) to \( C_k \)), then there is no need to perform the observation step for the corresponding belief factor \( b_k \). This is expressed formally in Theorem 2:

**Theorem 2.** If all \( x_t^{t+1} \in C_k \) are marginally independent of all \( y_j^{t+1} \in Y^{t+1} \) in \( \Delta_a \), then

\[
\forall s \in S : b_{k}^{t+1}(s_k) = \hat{b}_{k}^{t+1}(s_k).
\]

**Proof.** Proof in Appendix B. \( \square \)

Theorem 2 states that if the variables in \( C_k \) are independent of those in \( Y^{t+1} \), then the observation step for \( b_k \) can be skipped. However, even if \( C_k \) is not independent of \( Y^{t+1} \), it may be the case that the variables in \( C_k \) depend only on a subset \( Y_k \subset Y^{t+1} \) of the observation variables. Clearly, in such cases, it suffices to use \( Y_k \) rather than \( Y^{t+1} \) in the observation step. To account for this, we first note that the variables in \( Y^{t+1} \) may be correlated with each other. To preserve the correlations, we subdivide \( Y^{t+1} \) into clusters \( \hat{C}_l \subset Y^{t+1} \) and introduce the following assumptions:

\[
\begin{align*}
(A3) \ & \forall a : y_j^{t+1} \in \hat{C}_l \rightarrow (Y^{t+1} \cap pa_a(y_j^{t+1})) \subseteq \hat{C}_l \\
(A4) \ & \forall l \neq l' : \hat{C}_l \cap \hat{C}_{l'} = \emptyset
\end{align*}
\]

Assumptions (A3) and (A4) are analogous to (A1) and (A2), respectively, and essentially serve the same purposes for the observation step. To distinguish the clusters \( C_k \) and \( \hat{C}_l \), we
sometimes refer to the former as state cluster and to the latter as observation cluster. Given that (A3) and (A4) hold, we can redefine the observation step to

\[ b_k^{t+1}(s'_k) = \eta_2 \tilde{b}_k^{t+1}(s'_k) \prod_{l : \hat{C}_l \cap Y_k \neq \emptyset} \sum_{\hat{s} \in S(\overline{pa}_a^{t+1}(\hat{C}_l)) : \hat{s}_k = s'_k} \prod_{k' \neq k : \hat{C}_l \cap \overline{pa}_a^{t+1}(\hat{C}_l) \neq \emptyset} \tilde{b}_{k'}^{t+1}(\hat{s}_{k'}) \]  

(8)

where

\[ \Omega_l^{\alpha}(\hat{s}, o_{t+1}^{l+1}) = \prod_{y_j^{t+1} \in \hat{C}_l} P_a(y_j^{t+1} = (o_i^{t+1})_j | \overline{pa}_a(y_j^{t+1}) \leftrightarrow (\hat{s}, o_i^{t+1})) \]

and \( Y_k \subset Y^{t+1} \) is the set of observation variables which are not marginally independent of the variables in \( C_k \).

Given Theorem 2, one can see that (8) is equivalent to (7) if the observation variables are not clustered (or, equivalently, there is a single observation cluster \( \hat{C}_l = Y^{t+1} \)). However, it is important to note that if the observation variables are clustered (i.e. there are multiple observation clusters \( \hat{C}_l \)), then (8) is not necessarily equivalent to (7). To see this, it is helpful to compare the abstract formulations \( \prod_{j=1}^m \sum_{s} \Omega_s(o_j) b_s \) and \( \sum_{j=1}^m \prod_{s} \Omega_s(o_j) b_s \), where the former corresponds to (8) and the latter to (7). Therein, \( (o_1, ..., o_m) \in O \) is an observation, \( b_s \) is the probability of being in state \( s \in S \), and \( \Omega_s(o_j) \) is the probability of observing \( y_j = o_j \) in \( s \). These abstract formulations are equivalent for \( m = 1 \) or if \( b_s = 1 \) for some \( s \), but in all other cases they may not be equivalent. Nonetheless, if we fix the number of observation variables \( m \), then (8) approximates (7) closely as we increase the number of state variables \( n \). Our experiments indicate that it often suffices to use just a few more state variables than observation variables in order to obtain good approximations.

Finally, to show that it suffices to perform the observation step for \( b_k \) using only those clusters \( \hat{C}_l \) whose variables are not independent of the variables in \( C_k \), we observe that (8) is in fact a repeated application of (7) for every \( \hat{C}_l \), where the updated belief factor \( \tilde{b}_k^{t+1} \) is used in place of \( b_k^{t+1} \) in the subsequent application. Since every application has the same form as (7) (with \( Y^{t+1} = \hat{C}_l \)), we conclude that Theorem 2 holds, and hence the observation step can be skipped for clusters \( \hat{C}_l \) which are independent of \( C_k \).

5.4 Summary of PM

The preceding sections can be summarised as follows:

- **Representation**: The belief state \( b^t \) is represented as a product of \( K \) belief factors \( b^t_k \), such that \( b^t(s) = \prod_{k=1}^K b^t_k(s) \). Each belief factor \( b^t_k \) is a probability distribution over the set \( S(C_k) \), where \( C_k \subseteq X^{t+1} \) is a cluster of correlated state variables.

- **Transition step**: The transition step \( b^{t+1}_k \rightarrow \tilde{b}^{t+1}_k \) is performed using (6), for all clusters \( C_k \) which include active variables in \( \Delta^{a^t} \), or to which there is a causal path from an active variable in \( \Delta^{a^t} \). All other clusters are skipped.

- **Observation step**: The observation step \( \tilde{b}^{t+1}_k \rightarrow b^{t+1}_k \) is performed using (8), for all clusters \( C_k \) which are dependent on the observation variables \( Y^{t+1} \), using only those observation clusters \( \hat{C}_l \) which are relevant for \( C_k \). All other clusters are skipped.
Algorithm 3 provides a procedural specification of PM. The algorithm takes as inputs the action at time $t$, $a^t$, the subsequent observation at time $t+1$, $o^{t+1}$, and the belief factors at time $t$, $b^t_k$. As internal parameters, it uses a set of state clusters $C$, a set of observation clusters $\hat{C}$, and the set of DBNs $(\Delta^a)_{a \in A}$ which define the process. Lines 4 to 14 implement the transition step while lines 15 to 25 implement the observation step. Note that it suffices to execute lines 5 and 17 once in advance (or on demand) and to remember the results for future reference. The algorithm returns the updated belief factors $b^{t+1}_k$.

5.5 Space and Time Complexity

A belief factor $b_k$ has one element $b_k(s_k)$ for each $s_k \in S(C_k)$. Thus, the total space required to maintain $K$ belief factors $b_k$ is $\sum_{k=1}^{K} |S(C_k)|$. Furthermore, the size of the set $S(C_k)$ grows exponentially with the number of variables in $C_k$, hence the dominant growth factor in the space requirement is given by the largest cluster $C_k$ such that $|C_k| = \max_{k'} |C_{k'}|$. Therefore, we conclude that the space complexity of PM is in $O(\exp \max_k |C_k|)$, hence the representation is feasible for reasonably small clusters $C_k$.

Similarly, the number of operations required to perform the transition and observation steps is in the order of $2 \sum_{k=1}^{K} |S(C_k)|$ in the worst case (i.e. all clusters need to be updated in both steps). Specifically, line 11 and line 22 in Algorithm 3 are each executed once for every $s_k \in C_k$. The dominant growth factor is again given by the largest cluster $C_k$, hence the time complexity of PM is in $O(2 \exp \max_k |C_k|) = O(\exp \max_k |C_k|)$. Note that this assumes that the analysis performed by lines 5 and 17 in Algorithm 3 is done in advance.

The above time complexity is for the worst case in which all clusters need to be updated in the transition and observation steps. It is difficult to derive the time complexity for the average case because it is unclear what the average case is in terms of passivity. Even if we stipulate a certain average degree of passivity (e.g. 50% of all variables are passive), it would still be difficult to make a general statement about time requirements since this depends crucially on how the passive variables are distributed across the clusters. For example, even if a system has on average 90% passivity, if there is one active variable in each cluster then every cluster would need to be updated in the transition step. Thus, the only general statement we can make with regards to passivity is that the time complexity of PM can be refined to $O(\exp \max_{k \in C_T \cup C_O} |C_k|)$, where $C_T$ and $C_O$ include only those clusters that need to be updated in the transition and observation step, respectively.

5.6 Error Bounds

There are five possible sources of approximation errors in PM:

- If the clusters are correlated (i.e. (A1) or (A3) are violated)
- If the clusters are overlapping (i.e. (A2) or (A4) are violated)
- Generally in (8) if multiple observation clusters $\hat{C}_l$ are used

In the first two cases, the approximation error depends on the amount of correlation and overlap. If there is only little correlation and overlap between the clusters, then the
Algorithm 3 PM\((a^t, o^{t+1}, (b^t_k)_{C_k \in C}) \mid C, \hat{C}, (\Delta^a)_{a \in A}\)  

1: **Input:** action \(a^t\), observation \(o^{t+1}\), belief factors \((b^t_k)_{C_k \in C}\) 
2: **Parameters:** state clusters \(C\), observation clusters \(\hat{C}\), DBNs \((\Delta^a)_{a \in A}\) 
3: **Output:** updated belief factors \((b^{t+1}_k)_{C_k \in C}\) 
4: // **Transition step:**
5: \(C^* \leftarrow \text{SkippableClusters}(C, \Delta^a)\) 
6: for all \(C_k \in C\) do 
7: if \(C_k \in C^*\) then 
8: \(\hat{b}^{t+1}_k \leftarrow b^t_k\) 
9: else 
10: for all \(s'_k \in S(C_k)\) do 
11: \(\hat{b}^{t+1}_k(s'_k) \leftarrow \eta_1 \sum_{\bar{s} \in S(pa^t_{a^t}(C_k))} \prod_{k' : C_{k'} \cap pa^t_{a^t}(C_k) \neq \emptyset} b^t_{k'}(\bar{s})\) 
12: end for 
13: end if 
14: end for 
15: // **Observation step:**
16: for all \(C_k \in C\) do 
17: \(Y_k \leftarrow \left\{ y_j^{t+1} \in Y^{t+1} \mid \text{there is a directed path from } y_j^{t+1} \text{ to } C_k \in \Delta^{a^t} \right\}\) 
18: if \(Y_k = \emptyset\) then 
19: \(b^{t+1}_k \leftarrow \hat{b}^{t+1}_k\) 
20: else 
21: for all \(s'_k \in S(C_k)\) do 
22: \(\hat{b}^{t+1}_k(s'_k) \leftarrow \eta_2 \hat{b}^{t+1}_k(s'_k) \prod_{C_i \in \hat{C} : C_i \cap Y_k \neq \emptyset} \sum_{\bar{s} \in S(pa^t_{a^t}(\hat{C}_i)) : \bar{s}_k = s'_k} \prod_{k' \neq k : C_{k'} \cap pa^t_{a^t}(\hat{C}_i) \neq \emptyset} \hat{b}^{t+1}_{k'}(\bar{s})\) 
23: end for 
24: end if 
25: end for 
26: **return** \((b^{t+1}_k)_{C_k \in C}\) 

approximation error can be expected to be small. Conversely, if the clusters are strongly correlated and overlapping, then the approximation error can be expected to be large.
Theorem 3 (Boyen and Koller (1998)) work in Theorem 3, essentially states that the approximation error of PM (measured in terms of relative entropy) is bounded by the mixing rates of the process:

\[ KL(b||\tilde{b}) - KL(b'||\tilde{b}') \leq \epsilon^a. \]

The analysis also relies on the concept of mixing rates. Intuitively, the mixing rate \( \gamma^a \) of a DBN \( \Delta^a \) quantifies the degree of stochasticity of \( \Delta^a \). It depends on the mixing rates \( \gamma^a_k \) of the individual clusters \( C_k \):

Definition 8 (Relative entropy). Let \( \phi \) and \( \psi \) be two probability distributions defined over a set \( X \). The relative entropy from \( \phi \) to \( \psi \) is defined as

\[ KL(\phi||\psi) = \sum_{x \in X} \phi(x) \ln \frac{\phi(x)}{\psi(x)} \]

where \( \phi(x) > 0 \Rightarrow \psi(x) > 0 \).

Similar to Boyen and Koller (1998), we define the approximation error incurred by PM relative to the exact belief state. However, since we consider POMDPs with multiple actions \( a \in A \) (represented by the DBNs \( \Delta^a \)), we define the error for each action respectively:

Definition 9 (Approximation error). Let \( b \) be an exact belief state and \( \tilde{b} \) be the approximation by PM. After taking action \( a \), let \( b' \) be the exact update of \( b \) (using (1) and (2)) and \( \tilde{b}' \) be the PM-update of \( b \) (using (6) and (8)). Furthermore, let \( \tilde{b}' \) be the exact update of \( \tilde{b} \) (using (1) and (2)). We say that PM incurs error \( \epsilon^a \) in \( \Delta^a \) relative to \( b' \) if

\[ KL(b'||\tilde{b}') - KL(b'||\tilde{b}') \leq \epsilon^a. \]

Definition 10 (Mixing rate). The mixing rate of cluster \( C_k \subset X^{t+1} \) in \( \Delta^a \) is defined as

\[ \gamma^a_k = \min_{s',s''} \sum_s \min[T^a(s',s),T^a(s'',s)]. \]

If all \( C_k \) satisfy (A1) and (A2), and if all observation variables \( Y^{t+1} \) are in one observation cluster, then the mixing rate of \( \Delta^a \) is given by \( \gamma^a = (\min_k \gamma^a_k/r)^q \) where each cluster \( C_k \) depends on at most \( r \) and influences at most \( q \) other clusters \( C_{k'} \neq k \) (Boyen & Koller, 1998). In the worst case (that is, all (A1–A4) are violated), the minimal mixing rate is given by \( \gamma^a_k \) for the single cluster \( C_k = X^{t+1} \).

Finally, the main result of Boyen and Koller (1998), here restated in the context of our work in Theorem 3, essentially states that the approximation error of PM (measured in terms of relative entropy) is bounded by the mixing rates of the process:

Theorem 3 (Boyen and Koller (1998)). Let \( b' \) be an exact belief state and \( \tilde{b}' \) be the approximation by PM using clusters \( C_k \). Then, for any \( t \) with states \( s = (s^0, ..., s^t) \) and actions \( \bar{a} = (a^0, ..., a^{t-1}) \), we have

\[ E_{o^1, ..., o^t} [KL(b'||\tilde{b}')] \leq \frac{\max_{a \in \bar{a}} \epsilon^a}{\min_{a \in \bar{a}} \gamma^a} \]

where the expectation \( E \) is taken over all possible sequences of observations \( o^1, ..., o^t \) with probabilities \( P(o^1, ..., o^t) = \prod_{\tau=0}^{t-1} P^a(s^{\tau+1}, o^{\tau+1}), \) and where \( \epsilon^a \) and \( \gamma^a \) are defined as above.
6. Experimental Evaluation

We evaluate PM in two experimental domains: in Section 6.1 we evaluate PM in synthetic (i.e. randomly generated) processes with varying sizes and degrees of passivity, and in Section 6.2 we evaluate PM in a simulation of a multi-robot warehouse system. A brief summary of the experimental results is given in Section 6.3.

In our experience, two widely used monitoring methods are PF (Gordon et al., 1993) and BK (Boyen & Koller, 1998): PF, because it is very simple to implement and naturally lends itself to parallel execution; and BK, because it has firm theoretical error bounds and is readily available in many libraries, e.g. BNT (Murphy, 2001) and Infer.NET (Minka, Winn, Guiver, & Knowles, 2012). Therefore, we will compare PM against PF and BK in our experiments. See Section 2 for a more detailed discussion of these methods.

6.1 Evaluation in Synthetic Processes

We implemented the algorithms PF, BK, and PM in Matlab 7.13, where we used the Matlab toolbox BNT (Murphy, 2001) to implement BK. The algorithms were evaluated in synthetic processes of four different sizes, which are specified in Table 1.

<table>
<thead>
<tr>
<th>Process size</th>
<th># of x vars (n)</th>
<th># of y vars (m)</th>
<th># of states (#S)</th>
<th># of obs. (#O)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>10</td>
<td>3</td>
<td>&gt; one thousand</td>
<td>8</td>
</tr>
<tr>
<td>M</td>
<td>20</td>
<td>6</td>
<td>&gt; one million</td>
<td>64</td>
</tr>
<tr>
<td>L</td>
<td>30</td>
<td>9</td>
<td>&gt; one billion</td>
<td>512</td>
</tr>
<tr>
<td>XL</td>
<td>40</td>
<td>12</td>
<td>&gt; one trillion</td>
<td>4096</td>
</tr>
</tbody>
</table>

Table 1: Synthetic process sizes. All variables are binary.

The edges from $X^t / X^{t+1}$ to $X^{t+1}$ were sampled randomly using a mixture of Gaussians $G$, such that the probability of adding an edge $(x^t_i, x^{t+1}_j)$ or $(x^{t+1}_i, x^{t+1}_j)$ was specified by the combined density of $G$ at the value $j$. This was done so as to generate “areas” of correlated variables, which would then constitute natural candidates for state clusters $C_k$. The Gaussian mixtures were generated automatically for each process using Algorithm 4 in Appendix C. Furthermore, each $x^t_i$ was connected to at least one $x^{t+1}_j$, and each $x^{t+1}_j$ had at least one parent in $X^t$ or $X^{t+1}$. This was done to ensure that each variable had an effect in the process. Finally, the edges from $X^{t+1} / Y^{t+1}$ to $Y^{t+1}$ were sampled randomly with probability 0.1 such that each $y^{t+1}_j$ had at least one parent in $X^{t+1}$ or $Y^{t+1}$.

Each variable was binary and chosen to be passive with probability $p$, where we refer to $p$ as the degree of passivity. The variable distributions $P_{x_i}$ of $x^{t+1}_i \in X^{t+1}$ were generated randomly without bias, and the variable distributions $P_{y_j}$ of $y^{t+1}_j \in Y^{t+1}$ were generated with each probability sampled randomly from either [0.0, 0.2] or [0.8, 1.0] to obtain meaningful observations. Each process had two actions, $A = \{a_1, a_2\}$ (these were chosen randomly during the process), which were obtained by randomly choosing between 1 and 3 target variables whose distributions were resampled and edges from $X^t$ added with 0.1 probability, for each action $a \in A$. Each process started in a random initial state. All algorithms were tested on the same sequence of processes, initial states, chosen actions, and random numbers.
We used three different clustering methods, called $\langle \text{pc} \rangle$, $\langle \text{moral} \rangle$, and $\langle \text{modis} \rangle$. The methods were applied to the variables in $X_{t+1}$ without edges involving $X_t$ or $Y_{t+1}$:

- $\langle \text{pc} \rangle$ drops the directions of the edges (i.e. for any edge $x_{i,t+1} \rightarrow x_{j,t+1}$ it ads the reverse edge $x_{j,t+1} \rightarrow x_{i,t+1}$) and puts all variables between which there is a (undirected) path into one cluster. By definition, the resulting clusters satisfy all assumptions (A1–A4).

- $\langle \text{moral} \rangle$ connects all parents of a variable and drops the directions (it “moralises” the variables) and then extracts clusters of fully connected variables (“maximum cliques”). The resulting clusters may not satisfy any of the assumptions (A1–A4).

- $\langle \text{modis} \rangle$ is similar to $\langle \text{moral} \rangle$ but truncates the resulting clusters to make them disjoint (clusters are removed if they become a subset of another cluster). By definition, the resulting clusters satisfy (A2/A4), but not necessarily (A1/A3).

As an example, consider Figure 6 from Section 5.1. Here, $\langle \text{pc} \rangle$ would produce the cluster $C_1$ from Figure 6b, since all variables are connected by an undirected path. Furthermore, $\langle \text{moral} \rangle$ would produce the two clusters $C_1$ and $C_2$ from Figure 6c, which correspond to the two maximum cliques after moralising the variables in $X_{t+1}$. Finally, $\langle \text{modis} \rangle$ would produce the cluster $C_1$ from Figure 6c and the cluster $C_3$ from Figure 6a.

### 6.1.1 Accuracy

To compare the accuracy of the algorithms, we computed the relative entropy (cf. Definition 8) from exact belief states to the algorithms’ approximate belief states. Since exact belief states and relative entropy are hard to compute for large processes, we were able to do this for processes of size $S$ only. BK and PM used $\langle \text{pc} \rangle$, $\langle \text{modis} \rangle$, and $\langle \text{moral} \rangle$ clustering, where PM enforced (A1/A3) whenever necessary by modifying the variable distributions as described in Section 5.1. PM used the same clustering method to generate clusters of state variables ($C_k$) and observation variables ($\hat{C}_l$). PF was given the same amount of time needed by BK and PM when using $\langle \text{moral} \rangle$ clustering, for each process respectively. All algorithms were initialised with uniform belief states.

Figure 8 shows the relative entropy averaged over 1000 processes with 0%, 20%, 40%, 60%, 80%, and 100% passivity, respectively. The results show that PM $\langle \text{pc/modis} \rangle$ achieved a higher average accuracy than BK $\langle \text{pc/modis} \rangle$, and that PM $\langle \text{moral} \rangle$ achieved an average accuracy comparable to BK $\langle \text{moral} \rangle$. This indicates that violations of (A2/A4) introduce smaller errors than violations of (A1/A3). We also note that PM exhibited a significantly lower fluctuation than BK (with the exception of PM $\langle \text{moral} \rangle$ in 0% passivity). This can be a useful feature in the control task, since it means that PM was less surprised by the observations than BK. Finally, we note that the relative entropy of both methods was higher in processes with high passivity. This is explained by the fact that higher passivity implies higher determinacy, and hence lower mixing rates (cf. Definition 10), which are a crucial factor in the error bounds of PM and BK (cf. Theorem 3).

Note that PM $\langle \text{pc} \rangle$ did not produce exact belief states (i.e. zero relative entropy), despite the fact that the clusters generated by $\langle \text{pc} \rangle$ satisfy all assumptions (A1–A4). However, as mentioned in Section 5.6, and discussed in detail in Section 5.3, another possible source for approximation errors is if multiple observation clusters are used. Thus, the reason why PM $\langle \text{pc} \rangle$ did not produce exact belief states is that $\langle \text{pc} \rangle$ generated multiple observation clusters.
Figure 8: **Accuracy results.** Plots show relative entropy from exact to approximate belief states. Results are averaged over 1000 processes of size $S (n = 10, m = 3)$ where on average 0%–100% of non-target variables were passive. BK/PM used clustering methods $\langle pc \rangle$, $\langle moral \rangle$, and $\langle modis \rangle$. Results for PF not shown since relative entropy was too high.

We omit the results for PF since it produced a much higher relative entropy than BK and PM. This is because, given the times needed by BK $\langle moral \rangle$ and PM $\langle moral \rangle$, PF was only able to process between 100 and 300 state samples. Since each process has over 1000 states, this is not nearly enough to represent a uniform belief state. Moreover, the fact that the processes have very high stochasticity (thus, each state has relatively high probability)
Figure 9: **Timing results.** (a–d) Plots show number of seconds required for 1000 transitions on a dual-core machine with 2.4 GHz. Times were averaged over 1000 processes of size S, M, L, and XL, respectively. Passivity of $p\%$ means that on average $p\%$ of non-target variables were passive. BK and PM used $\langle \text{moral} \rangle$ clustering. PF was optimised for binary variables and used number of samples to achieve accuracy of BK and PM, respectively. PM was run with 1 (PM-1), 2 (PM-2), 4 (PM-4) parallel processes. (e) Average percentage of clusters which were updated in the transition and observation steps.

means that PF would require significantly more state samples to achieve the same accuracy as BK and PF, as shown in the next section.

### 6.1.2 Timing

We measured computation times in processes of sizes S, M, L, XL with passivities of 25%, 50%, 75%, 100%, respectively, on a dual-core machine with 2.4 GHz. BK and PM used $\langle \text{moral} \rangle$ clustering, which seemed most appropriate for a fair comparison since it produced consistently similar accuracy for both monitoring algorithms. PF performed a heuristic search in every process to find the right number of samples to achieve an accuracy which on average was as good as BK and PM, respectively, in the final 20% of the process. As this involved computing exact belief states and relative entropies, we were able to use PF in
processes of size $S$ only. We tested PM with 1, 2, and 4 parallel processes, which in every transition were randomly allocated the same number of clusters.

Figures 9a – 9d show the times for 1000 transitions (averaged over 1000 processes) and Figure 9e shows the average percentage of clusters updated in the transition and observation steps of PM. The times for PM include the times needed to find skippable clusters in the transition and observation steps, which was done once in advance for each action. The results show that PM was able to minimise the time requirements significantly by exploiting passivity. First, we note that there were only marginal gains from 25% to 50% passivity, despite the fact that PM updated 14% fewer clusters in the transition step. This is because these clusters were mostly very small. However, there were significant gains from 50% to 75% passivity with average speed-ups of 11% (S), 14% (M), 15% (L), 18% (XL), and from 75% to 100% passivity with further average speed-ups of 11% (S), 33% (M), 46% (L), 49% (XL). This shows that the computational gains can grow significantly with both the degree of passivity and the size of the process.

Our results show that PM consistently outperformed BK in all process sizes. However, while the times of both algorithms grew exponentially in the size of the process (with factors of approximately 2 for BK and 2.5 for PM), the relative difference between PM and BK was decreasing. We attribute this to our naive implementation of PM, which was essentially a straightforward implementation of the equations given in Section 5. We expect that large improvements can be made with a more sophisticated implementation.

How were BK and PF affected by passivity? As expected, the performance of BK was nearly unaffected by the varying degrees of passivity. The junction tree algorithm used in BK benefited marginally from an increased sparsity in the process, but the computational gains were minimal. We were at first unable to use PF as it required too many samples (between 10k and 200k) to achieve the same accuracy as BK/PM. This is due to the very high variance in the processes, which is a known problem of PF (Gordon et al., 1993). In order to investigate the effect of passivity on PF, we implemented a version of PF which was strictly optimised for binary variables. Interestingly, we found that passivity had an adverse effect on the performance of PF, requiring it to use exponentially more samples with increased passivity (see Figure 9a). This makes sense if we view PF as a factored approximation method (such as PM and BK) which means that Theorem 3 applies. However, because PF puts all variables into a single cluster (since it is not actually a factored method), the mixing rate of the process will be much lower than for PM and BK (as discussed in Section 5.6) and, thus, the error bounds are less tight. To compensate this, PF requires significantly more samples for increased passivity.

6.2 Evaluation in Real-world System

In this section, we demonstrate how passivity can occur naturally in a real-world system and how PM can exploit this to accelerate the monitoring task. To this end, we consider a multi-robot warehouse system in which the robots’ task is to transport goods within the warehouse. Specifically, we simulate a warehouse in the style of KIVA (Wurman, D’Andrea, & Mountz, 2008), which is a large commercial multi-robot warehouse system used by global companies such as Staples and Amazon (D’Andrea, 2012). POMDPs are a useful framework
to model warehouse systems and have been used in several related works, e.g. (Hariharan & Bukkapatnam, 2009; Estanjini, Lin, Li, Guo, & Paschalidis, 2011).

6.2.1 Specification of the Warehouse System

Figure 10 shows the initial state of the warehouse simulation. The warehouse consists of 2 workstations (W1, W2), 4 robots (R1–R4), and 16 inventory pods (I1–I16). Each robot can move forward and backward, turn left and right, load and unload an inventory pod (if positioned under the pod), or do nothing. As in KIVA, robots can move under inventory pods unless they are carrying a pod, in which case the other pods become obstacles. The move and turn operations are stochastic in that the robot may move/turn too far (3% chance) or do nothing (2% chance). Each robot possesses two sensors, one telling it which inventory pod it has loaded (if any) and one for the direction it is facing. The direction sensor is noisy in that a random direction may be reported (3% chance). Each robot maintains a list of tasks in the form of “Bring inventory pod I to workstation W” (yellow area around W) and “Bring inventory pod I to position (x,y)”. How these tasks are executed depends on the control mode, of which we use two in our simulations:

11. Centralised mode — A central controller maintains a belief state \( b^t \) about the warehouse system. At each time \( t \), it samples 100 states from \( b^t \) and removes all duplicate states, resulting in the set \( \hat{S} = \{ \hat{s}_1, \hat{s}_2, \ldots \} \). It then resamples a state \( \hat{s}^* \in \hat{S} \) with probabilities \( w(\hat{s}^*) = b^t(\hat{s}^*)/\sum_q b^t(\hat{s}_q) \). Based on \( \hat{s}^* \) and the current task of each robot, it performs an \( A^* \) search (Hart, Nilsson, & Raphael, 1968) (with Manhattan distance) in the space of joint actions to find the optimal action for each robot. After executing their actions, the robots send their sensor readings to the controller, and the controller updates its belief state using the sensor readings.

11. Our control modes are ad hoc and often make suboptimal decisions. However, we found that current solution techniques for (DEC-)POMDPs, including point-based methods, were infeasible in this setting. Nonetheless, the quality of the decisions made by our control modes largely depends on the accuracy of the belief states, hence it is important that the belief states are updated accurately. Therefore, the control modes were sufficient for our purposes.
**Autonomous mode** — Each robot maintains its own belief state and there is no communication between the robots. The only knowledge the robots have about each other are their current tasks, communicated by the task allocation module. At each time $t$, each robot samples the set $\hat{S}$ and state $\hat{s}^*$ as is done in the centralised mode. Treating the other robots as static obstacles, it performs an $A^*$ search based on $\hat{s}^*$ and its current task to find an action $a^t$. This is repeated for each other robot $r$ in all states $\hat{s}_q \in \hat{S}$, resulting in actions $a_{r,q}$ which are used to obtain distributions $\pi_r : A \rightarrow [0, 1]$ ($A$ is the set of all actions) with $\pi_r(a) = \sum_{q : a_{r,q}=a} w(\hat{s}_q)$. The robot then executes its action $a^t$ and updates its belief state using its sensor readings and the distributions $\pi_r$ to average over the other robots’ actions.

The tasks are generated by an external scheduler in time intervals sampled from $U[1, 10]$. Each generated task is assigned to one of the robots through a sequential auction (Dias, Zlot, Kalra, & Stentz, 2006), in which the robots’ bids are computed as their total number of steps needed to solve all of their current tasks and the auctioned task, averaged over all states in $\hat{S}$. The robot with the lowest bid gets the task.

6.2.2 Warehouse DBN, Clusters & Passivity

Figure 11 shows an example DBN for a smaller warehouse with one inventory pod and two robots. Each inventory pod $I$ is represented by two variables, $I.x$ and $I.y$, which correspond to the x and y position of the inventory pod. Each robot $R$ is represented by four variables: $R.x/R.y$ for its x/y position, $R.d$ for its direction, and $R.s$ for its status. The status of a robot $R$ is either $R.s=0$ (unloaded) or $R.s=I$ (loaded with inventory pod $I$). Constants such as the size of the warehouse and the positions of the workstations are omitted in the DBN. There are four types of clusters: The I-clusters (C1–C4) preserve the correlation that if $R$ is loaded with $I$, then $I$ must always have the same position as $R$ (there are two I-clusters for each $(I,R)$ pair); The R-clusters (C5) and S-clusters (C6), respectively, preserve the correlation that no two robots can have the same position or carry the same inventory pod (there is one R/S-cluster for each $(R_a,R_b)$ pair with $a > b$); And, finally, the D-clusters (C7, C8). PM uses singleton observation clusters (i.e. there is one cluster for each observation variable).

There are some differences between the DBNs for the centralised and autonomous modes (Figure 11 uses the centralised mode). In the centralised mode, there is one DBN for each action combination of the robots. Since the controller observes all $R.s$ noise-free, it can add edges from $R.x/R.y$ to $I.x/I.y$ if $R.s=I$ or remove them otherwise to simplify the inference (thus, in Figure 11, $R1$ is loaded with $I1$ and $R2$ is unloaded). In the autonomous mode, each robot only observes its own sensor readings, hence it can add or remove edges only for itself, while edges for all other robots must be permanently added. This also means that the other robots’ status variables ($R.s$) must be linked to all $I.x/I.y$ and, therefore, included in the I-clusters (to preserve the correlation that I must have the same position as $R$ if $R$ is loaded with $I$). Moreover, since each robot only knows its own action, there is one DBN for each of its own actions, and all variables associated with the other robots are active (the distributions $\pi_r$ defined in the previous section are used to average over their actions).

How many clusters PM updates in the transition step depends on the actions of the robots. For instance, in the centralised mode DBN shown in Figure 11, $R1$ is moving and $R2$ is turning. This means that $R1.x$, $R1.y$, and $R2.d$ are active variables (dashed circles)
Exploiting Causality for Efficient Monitoring in POMDPs

Figure 11: Example DBN of a smaller warehouse system consisting of one inventory pod (I1) and two robots (R1, R2). The DBN implements the joint action in which R1 moves and R2 turns. Solid white circles mark passive variables, dashed white circles mark active variables, and solid grey circles mark observation variables. The coloured areas represent the state clusters \( C_1 \) to \( C_8 \).

while all other variables are passive, corresponding to a passivity of 70%. In this DBN, PM updates the clusters C1, C2, C5, and C8, since they each contain active variables, and it also updates the clusters C3 and C4, since there are directed paths from active variables (R1.x and R1.y) to each of them. Therefore, the only clusters which are not updated are C6 and C7. Although this saving may appear somewhat small, recall that our focus and hypothesis pertain to how passivity can be exploited to a great effect in large processes. The full warehouse in our experiment contains 16 inventory pods and 4 robots, resulting in 48 variables with 128 I-clusters, 6 R-clusters, 6 S-clusters, and 4 D-clusters. If we consider a similar situation in this warehouse, in which one robot moves with an inventory pod while the others turn, then the savings can be much greater. For example, say R1–3 turn while R4 moves with inventory pod I1. Then, PM updates only 3 of 6 R-clusters (those containing R4), 0 of 6 S-clusters (since no status change), 3 of 4 D-clusters (for R1–3), and 38 of 128 I-clusters (32 I-clusters containing R4 plus 6 I-clusters from R1–3 for I1), amounting to a total saving of 69.44% of clusters which do not need to be updated.

6.2.3 Results

We implemented PF, BK, and PM in C#, using the framework Infer.NET (Minka et al., 2012) to implement BK. This allowed BK to exploit sparsity in the process and offered improved memory handling. PM was optimised for sparsity in (6) and (8), respectively, by summing over states \( \bar{s} \) for which all \( b_{t-1}^{t+1} \) are positive. PF naturally benefits from sparsity as it allows it to concentrate the samples on fewer states. The number of samples used in PF was set in such a way that the controller decisions were invariant of the random numbers used in the sampling process of PF. This was done to ensure that the results were repeatable.
Finally, to maintain sparsity in the process, each probability in the belief states lower than 0.01 was set to 0.

Figure 12 shows the time per transition averaged over 20 different runs with 100 transitions each (measured on a dual-core machine with 2.4 GHz). The times for PM include the time needed to find skippable clusters for the transition and observation steps, which was done once on demand for every previously unseen DBN. Each algorithm started with an exact initial belief state, shown in Figure 10. In the centralised mode, PM was able to outperform BK on average by 48% and PF by 36%, which needed 20,000 samples to produce consistent (i.e. repeatable) results. In the autonomous mode, PM outperformed BK on average by 17% and PF by 32%, which now needed 45,000 samples to produce consistent results, due to the increased stochasticity in the process. All differences were statistically significant, based on paired t-tests with 5% significance level. Note that BK and PM were slower in the autonomous mode since the corresponding DBNs had much higher inter-connectivity. Also, PM updated more clusters since there were more active variables.

The number of states in the warehouse system (including invalid states) exceeded $10^{45}$ states. Therefore, we were unable to compare the accuracy of the tested algorithms in terms of relative entropy. Instead, we compared their accuracy based on the results of the task auctions and the number of completed tasks by the end of each run. This gives a good indication of the algorithms’ accuracy since both the outcome of the auction and the number of completed tasks depend on the accuracy of the belief states. In the centralised mode, the algorithms generated over 95% identical task auctions and completed 15.7 (BK), 15.5 (PM), and 15.2 (PF) tasks on average. In the autonomous mode, they generated over 93% identical auctions and completed 12.1 (BK), 12.2 (PM), and 11.7 (PF) tasks on average. In both modes, none of these differences were statistically significant.

6.3 Summary of Results

The experimental results show that PM produces belief states with competitive accuracy: In the synthetic processes, PM achieved an accuracy which on average was better or comparable to the accuracy of the alternative methods. In the warehouse system, PM was able to
complete a statistically equivalent number of tasks as compared to the other methods, which means that its accuracy must have been equivalent or at least comparable.

Furthermore, the experimental results show that PM executed the belief updates significantly faster than the alternative methods: In the synthetic processes, PM (using no parallel processes) outperformed BK by up to a factor of 2.8 in the largest process (XL), while PF took too much time to achieve an accuracy comparable to PM. In particular, the results show that the computational gains can grow significantly with both the degree of passivity and the size of the process. In the warehouse system, PM outperformed the alternative methods by up to 48%, which is a significant saving considering the size of the state space (more than $10^{45}$ states). Furthermore, the computational gains where much higher in the centralised control mode than in the autonomous control mode, where the latter had a significantly lower degree of passivity. Therefore, this again shows that high degrees of passivity can bear great potential for the monitoring task.

7. No Free Lunch for PM

For reasons given in Sections 1 and 2, our view is that no monitoring method is generally best for all types of systems. Instead, each method explicitly or implicitly assumes a certain structure in the system which it attempts to exploit in order to render the monitoring task more tractable. Typically, the methods are tailored in such a way with respect to this structure that they perform well if the structure is present in the system, but suffer a significant loss in performance if the structure is absent. Therefore, the exploited structure essentially defines the strengths and weaknesses of a monitoring method.

A formal account of this view is given by the “No Free Lunch” theorems (Wolpert & Macready, 1995, 1997) which state that, intuitively speaking, any two algorithms have equivalent performance when averaged over all possible instances of the problem. In other words, if there are classes of problem instances for which algorithm A has better performance than algorithm B, then there must be other classes of problem instances for which A has worse performance than B. Then, the question is: for what class of problem instances (that is, systems) can PM be expected to achieve good performance? This class is essentially described by the following three criteria:

**Degree of passivity** — PM attempts to accelerate the monitoring task by omitting the belief update for as many state clusters ($C_k$) as possible. In the transition step, this depends on the passivity of the variables in the clusters. In the ideal case, the system exhibits a high degree of passivity such that PM can omit the transition step for many state clusters. In the worst case, the system has no passive variables at all, and PM has to update all belief factors in the transition step.

However, as discussed in Section 5.5, a high degree of passivity is not necessarily sufficient to infer that many clusters can be skipped in the transition step, since the passive variables could be distributed in such a way that no cluster can be skipped (e.g. if the passive variables are distributed uniformly amongst the state clusters). Therefore, in an optimal case, the passivity is concentrated on correlated state variables such that passive variables end up in the same clusters.
Size of state clusters — As discussed in Section 5.5, the space and time complexity of the belief state representation in PM is exponential in the size of the largest state cluster. Therefore, in the ideal case, the relevant variable correlations can be captured in small state clusters and the cost of storing the belief factors and performing the update procedures is little. In the worst case, large state clusters are required to retain the variable correlations and the cost of storing and updating belief factors is large.

Besides this complexity issue, another reason why the state clusters should be small is because of the way in which PM performs the transition step. One pre-requisite for omitting the transition step for a cluster is that all variables in the cluster are passive. If there are many variables in one cluster, then it is less likely that all variables in the cluster are passive, and, hence, it is less likely that the cluster can be skipped.

Structure of observations — A third criterion, though arguably less important than the other criteria, is the structure of the observations (i.e. the way in which the observation variables depend on the state variables) and the size of the observation clusters ($\hat{C}_l$). PM attempts to accelerate the observation step by skipping all those state clusters whose variables are structurally independent of the observation, and, if a cluster cannot be skipped, by incorporating only those observation clusters which are relevant to the update. Therefore, in the ideal case, only a fraction of the state clusters depend on the observation, and the relevant correlations between observation variables can be captured in small observation clusters. In the worst case, all state clusters depend on the observation in some sense, and the structure of the observation does not allow for an efficient clustering.

Thus, in summary, PM is most suitable for systems with high degrees of passivity, in which the relevant variable correlations can be captured in small state and observation clusters. On the other hand, PM may not be suitable if there is no or only low degrees of passivity, and if large state and observation clusters are necessary to retain the relevant variable correlations in the system.

In addition to identifying the class of systems for which a monitoring method is suitable, it is also important to justify the practical relevance of this class. In this work, we are primarily interested in robotic and other physical systems. Such systems typically exhibit a number of features: First of all, robotic systems usually have some causal structure, e.g. (Pearl, 2000; Mainzer, 2010). Passivity, as a specific type of causality, can be observed in many robotic systems, including the robot arm used in our examples and the multi-robot warehouse system in Section 6.2. Furthermore, robotic systems most typically have a modular structure, in which each module is responsible for a specific subtask and may interact with other modules. This modular structure often allows for an efficient clustering, in the sense that each module corresponds to a cluster of correlated system variables. Finally, the sensors used in robotic systems typically only provide information about certain aspects of the system, and some components of the system may not benefit from some of the sensor information. These features correspond to the criteria (above) which specify the class of systems for which PM is a suitable monitoring method. Therefore, we believe that this class is practically justified.
8. Conclusion

This paper explores the idea of exploiting causality in an automatic way to accelerate the monitoring task in POMDPs. To this end, we consider a specific type of causality, called *passivity*, which essentially pertains to how variables cause changes in other variables. In order to demonstrate the potential of exploiting passivity, we develop a novel monitoring method, called *Passivity-based Monitoring* (PM), which uses a factored belief state representation and exploits passivity to perform selective belief updates. PM produces exact belief states if certain assumptions are met and approximate belief states otherwise, where the approximation error is bounded by the degree of uncertainty in the process. We have shown empirically that PM outperforms two standard methods, PF and BK, in synthetic processes with varying sizes and degrees of passivity, and in an example of a real-world system with different degrees of passivity. Our results show that the computational gains can grow significantly with both the size of the process and the degree of passivity, while achieving competitive accuracy. Therefore, this supports our hypothesis that factored beliefs and passivity can indeed be a key combination in large systems.

It is important to stress that the primary purpose of this work is not to propose yet another monitoring method. Rather, the purpose of this work is to show that if a system exhibits a high degree of causal structure, then there can be great potential in exploiting this structure to render the monitoring task more tractable, as demonstrated clearly in our experiments. Therefore (as discussed in detail in Section 7), this work is relevant for complex real-world systems with high degrees of causality, including robots used in homes, offices, and industrial factories, where the monitoring task may constitute a major impediment due to the often very large state space of the system.

8.1 Extensions

There are several interesting directions for future work:

**Alternative types of causality** — As discussed earlier in this paper, passivity is merely one example of a causal relation, and other types of causal relations can be identified (Pearl, 2000). An interesting direction for future work would be to formulate alternative types of causal relations which can be exploited in ways similar to how PM exploits passivity, or perhaps in ways other than that. Alternatively, it would be interesting to see if the definition of passivity could be relaxed such that more variables fall under this definition, and such that the principal idea behind PM is still applicable.

**Efficient implementation of PM** — We used a “proof of concept” implementation of PM in our experiments, which means that the equations which define PM were implemented directly (i.e. as given). While the results of our experiments confirm that causality can be exploited to a great effect, we believe that the performance of PM could be further improved with a more sophisticated implementation. For example, it would be interesting to see if PM can be implemented efficiently on a graph structure, such as BK. In this regard, it is important to note that PM keeps the belief factors separate during the update procedure in order to perform selective updates. This is in contrast to BK, which combines the belief factors in the update procedure (more precisely, the junction tree algorithm (Lauritzen & Spiegelhalter, 1988) used in BK requires this)
and separates them afterwards in the “projection step” (Boyen & Koller, 1998). Thus, implementing PM on a graph structure is not a trivial task.

Causality in existing methods — As discussed in Section 2, existing monitoring methods do not exploit causal structure in systems, or at least not explicitly and effectively. It would be interesting to know to what extend existing monitoring methods could be modified to account for causal structure in systems. For example, Ng et al. (2002) proposed a method which combines the concepts of PF and BK. Perhaps similar combinations are possible with PM and other existing methods.

Appendix A. Proof of Theorem 1

To proof Theorem 1, it will be useful to first establish the following lemma:

**Lemma 1.** If (A1) holds and all $x_{i}^{t+1} \in C_k$ are passive in $\Delta^a$, then

$$\forall s, s': T_k^a(s, s'_k) = 1 \iff s_k = s'_k.$$  

**Proof.**

$\Rightarrow$: The fact that (A1) means that $\Phi_{a,i} \subseteq C_k$ for all $x_{i}^{t+1} \in C_k$. Since all $x_{i}^{t+1} \in C_k$ are passive in $\Delta^a$, it follows that all $x_{j}^{t} \in \Phi_{a,i}$ are passive in $\Delta^a$, for all $\Phi_{a,i}$. Therefore, given $T_k^a(s, s'_k) = 1$ and clause (ii) in Definition 4, it must follow that $s_k = s'_k$.

$\Leftarrow$: Follows directly by (A1) and the fact that all $x_{i}^{t+1} \in C_k$ are passive in $\Delta^a$.

Using Lemma 1, we can give a compact proof of Theorem 1:

**Theorem 2.** If (A1) and (A2) hold, and if all $x_{i}^{t+1} \in C_k$ are passive in $\Delta^{a'}$, then

$$\forall s : \hat{b}_{k}^{t+1}(s_k) = b_{k}^{t}(s_k).$$  

**Proof.**

$$\hat{b}_{k}^{t+1}(s_k) = \eta_1 \sum_{\bar{s} \in S(\text{pa}^t_{a,t}(C_k))} T_{k}^{a'}(\bar{s}, s'_k) \prod_{k': C_{k'} \cap \text{pa}^t_{a,t}(C_k) \neq \emptyset} b_{k'}^{t}(\bar{s}_{k'})$$

$$= \eta_1 \sum_{\bar{s} \in S(\text{pa}^t_{a,t}(C_k)): s_k = s'_k} \prod_{k': C_{k'} \cap \text{pa}^t_{a,t}(C_k) \neq \emptyset} b_{k'}^{t}(\bar{s}_{k'})$$

$$= \eta_1 b_{k}^{t}(s_k) \sum_{\bar{s} \in S(\text{pa}^t_{a,t}(C_k)): s_k = s'_k} \prod_{k': C_{k'} \cap \text{pa}^t_{a,t}(C_k) \neq \emptyset} b_{k'}^{t}(\bar{s}_{k'})$$

$$= \eta_1 b_{k}^{t}(s_k)$$

$$= b_{k}^{t}(s_k). \quad (\eta_1 = 1 \text{ since } b_{k}^{t} \text{ normalised})$$


Appendix B. Proof of Theorem 2

To proof Theorem 2, we first note the following proposition:

**Proposition 1.** If all $x^{t+1}_i \in C_k$ are marginally independent of all $y^{t+1}_j \in Y^{t+1}$ in $\Delta^a$, then

$$\forall s, s': (\land_{k' \neq k} s_{k'} = s_{k'}) \rightarrow \Omega^a(s, o^t) = \Omega^a(s', o^t).$$

This proposition follows directly by definition.

Using Proposition 1, we can give a compact proof of Theorem 2:

**Theorem 2.** If all $x^{t+1}_i \in C_k$ are marginally independent of all $y^{t+1}_j \in Y^{t+1}$ in $\Delta^a$, then

$$\forall s : b^{t+1}_k(s_k) = \hat{b}^{t+1}_k(s_k).$$

**Proof.**

$$b^{t+1}_k(s'_k) = \eta_2 \hat{b}^{t+1}_k(s'_k) \sum \Omega^a(\tilde{s}, o^{t+1}) \prod_{k' \neq k, k' \cap \text{pa}^{t+1}_a(Y^{t+1}) \neq \emptyset} \hat{b}^{t+1}_{k'}(\tilde{s}_{k'}) \quad \text{Prop1} \quad \text{constant } \alpha, \text{ independent of } s'_k$$

$$= \frac{\hat{b}^{t+1}_k(s'_k) \alpha}{\sum s''_k \hat{b}^{t+1}_k(s''_k) \alpha}$$

$$= \frac{\hat{b}^{t+1}_k(s'_k)}{\sum s''_k \hat{b}^{t+1}_k(s''_k)}$$

$$= \hat{b}^{t+1}_k(s'_k).$$
Appendix C. Mixture of Gaussians

Algorithm 4 MixtureOfGaussians(n)

1: Input: number of system states n
2: Parameters: $\lambda \leftarrow 4$, $\sigma_{\text{min}} \leftarrow \frac{5}{10}$, $\sigma_{\text{max}} \leftarrow \frac{n}{10}$
3: Output: mixture of Gaussians $G$
4: $G \leftarrow \emptyset$
5: $R \leftarrow \{(1, ..., n)\}$
6: while $R \neq \emptyset$ do
7: $R \leftarrow$ some element of $R$
8: $R \leftarrow R \setminus \{R\}$
9: $\mu \leftarrow R([\text{rand} \ast |R|])$ // rand returns random number from $(0, 1)$
10: $\sigma \leftarrow \min\left[\sigma_{\text{max}}, \max\left[\sigma_{\text{min}}, \text{rand} \ast \frac{\min[\mu - R(1), \mu + R(|R|)]}{\lambda}\right]\right]$ // Mean and variance of Gaussian
11: $G \leftarrow G \cup \{(\mu, \sigma^2)\}$
12: $R_- \leftarrow (R(1), R(2), ..., R(p))$ such that $R(p) < \mu - \sigma\lambda$
13: $R_+ \leftarrow (R(q), R(q + 1), ..., R(|R|))$ such that $R(q) > \mu + \sigma\lambda$
14: if $R_- \neq \emptyset$ then
15: $R = R \cup \{R_-\}$
16: end if
17: if $R_+ \neq \emptyset$ then
18: $R = R \cup \{R_+\}$
19: end if
20: end while
21: return $G$

References


