On Robust Markov Analysis for Reliability Assessment of Complex Systems using Imprecise Markov Chains

Daniel Krpelačk 1,2, Frank P.A. Coolen 1 and Louis J.M. Aslett 1

1 Department of Mathematical Sciences, Durham University, Durham, UK
2 Department of Applied Mathematics, VŠB-Technical University of Ostrava, Ostrava, CZ

Abstract—Markov analysis is a wide-spread tool for modelling interactions among components in complex systems. It is based on modelling the evolution of system’s component states by Markov Chains. But, as in many other uncertainty models, it might often be overly optimistic to assume that we can construct a precise stochastic model which properly captures the uncertainties present in the investigated system. This issue is addressed by the theory of Imprecise Probabilities and, specifically for stochastic processes, by the theory of Imprecise Markov Chains. In this paper, we will demonstrate how Imprecise Markov Chains can not only serve as a robust alternative to classical stochastic models, but also how they can facilitate analyses by the means of problem dimension reduction and also, by the means of deliberate model construction, enable analyses which would not be possible by using solely precise probability models.

Keywords: System Reliability, Imprecise Markov Chains, Survival Signatures, Robust Bayesian Inference.

I. INTRODUCTION

Modelling reliability of a complex systems is a daring task. One of the difficulties lies in properly capturing the interactions among the system components during the system’s functional lifetime. In order to model these complex interactions, it is necessary to use flexible models to capture the dynamical evolution of the system component states. Markov analysis is used for this purpose. The stochastic evolution is modelled by (Continuous Time) Markov Chains [9, 10]. But the introduced complexity poses a challenge to properly capture all the possible interactions. We will therefore demonstrate the use of a framework for statistical inference of the Markov Models in order to ground the models by empirical evidence, and to robustify them by using models from the theory of Imprecise Probability (IP; [1, 2]). One way of viewing the IP theory, is from the point of view of sensitivity analysis [11]. Instead of requesting to specify a single stochastic model, or a single sampling distribution, or a single prior distribution for Bayesian inference, we seek ways to analyse a whole set of them at once. Once working with a set of models, we lose the precision of the assertions we are used to from classical uncertainty analysis. Instead, we can infer bounds on the assertions of interest, guaranteed to include the correct answer (as long as the correct model is included in the set of assumed laws). The general aim of inference in the IP theory is to assess lower and upper bounds on functionals of interest (probabilities of events and expected values). For a set of probability measures \( \mathbb{P} \) and some random variable \( Y \), we define the lower and upper expectations as

\[
E_L Y := \inf_{P \in \mathbb{P}} E_P Y, \quad E_U Y := \sup_{P \in \mathbb{P}} E_P Y. \quad (1)
\]

The analogies for lower and upper probabilities follow from the representation of events by indicator functions.

Due to construction of the IP theory, we only need to focus on either lower or upper expectation since, for so-called coherent models, one can be derived from the other via conjugacy relation \( \mathbb{E}[Y] = -\mathbb{E}[-Y] \).

In this paper, we will focus on an application of results from a sub-field of IP theory, on theory of Imprecise Markov Chains [7]. Recent development concluded in efficient computational schemes which enable us to conduct robust inferences on stochastic dynamical systems [8]. An outline of the numerical methods will be presented in Section [V]. It has also been recently discovered how to perform Robust Bayesian Inference on the characteristics of the process dynamics [8], as will be shown in Section [VI]. Imprecise Markov Chains can also be used to reduce complexity of high-dimensional problems by the technique of lumping. This is not unknown to precise Markov Chain models, but its imprecise variant provides more flexibility because the lumping generally results in an Imprecise Markov Chain model [4]. This will be shown in Section [VII-B].

Our selected application field is Reliability Theory. We will combine the above-mentioned advancements in the field of Imprecise Markov Chains with modelling techniques of Reliability Theory. The state-space lumping can be naturally combined with Survival Signatures, which also serve as a dimension reduction technique [1]. These will be covered in Section [VII-A] and in Section [VII] in general. We will further use the flexibility of the IP theory in order to propose a novel method to 1) enable statistical inference for Markov Reliability analysis for realistic scenarios at all, and 2) include additional
high-level information into the models in order to decrease the imprecision of the answers. This will be explained in Section III and demonstrated on an example of reliability inference from a collection of observations for a simple system in Section IX.

II. NOTATION

We will refer to binary random variables and random events interchangeably since there exists a bijection between these representations (an event obtains if the random variable equals 1). Which of them is used should be evident from the context. Random variables will be denoted by capital letters and their realisations by a lower-case one, as is usual.

III. SYSTEM RELIABILITY

Random variable $X_i$ will represent the event that component $i$ is functioning (the event obtains iff $X_i = 1$). Suppose we have a system composed of $N$ binary components and denote $X = \{0, 1\}^N$ the joint state-space of the system components. We will denote by $X$ the random vector of component states. The relation between the component states and the system state can be modelled by a structure function $\varphi : X \rightarrow \{0, 1\}$ such that $\varphi(x) = 1$ iff $x$ results into a functioning system state. As usual, we will denote $(\Omega, \mathcal{A}, P)$ the underlying probability space and the induced probability measures of random variables by subscripting $P$.

Reliability of a system is defined as the probability that the system is functioning, thus

$$Rel = P(\varphi(X)) = E[\varphi(X)]$$

(2)

$$= \sum_{x \in \{0, 1\}^N} \varphi(x)P_X(x).$$

(3)

We will further assume that the investigated systems are coherent, which implies that $\varphi$ is a non-decreasing function of $x$.

Many systems are engineered such that they are functional at the instant of deployment but their components will deteriorate over time. In such scenarios, we investigate how the system reliability changes in time. The component states are here modelled as stochastic processes $X_i(t)$ and the joint states as $X(t)$ for some $t \geq 0$. The simplest case is when all the $X_i$ are non-increasing processes (once a component fails it will not recover its functionality). Then we can equivalently model just the time to the component’s failure, say $T_i$, and define the stochastic process as $X_i(t) = \{T_i > t\}$.

The dynamics of the components’ states will induce a stochastic process on the state of the system. Two basic assessments about the system are then usually of interest, the system availability $A(t) = E[\varphi(X(t))]$ (the probability that a system is operational at a given time instant, regardless of its history) and the system reliability $Rel(t) = P(\forall \tau \leq t : \varphi(X(\tau)) = 1)$, also called the survival function (the probability that the system has not failed before the enquired time instant). These two assessments are equivalent if the system is coherent and the component state processes are non-increasing.

The evolution of component states may take complex form and include interaction among the components states and their deterioration rates. We will model the joint component state processes by Markov Chains on the joint component state-space since they can capture various interactions among component state dynamics (load-share system, spare components, component renewals, etc).

IV. MARKOV ANALYSIS

In this section, we will focus on the Markov method for modelling the time evolution of component states. We will revise some of the basic elements of Markov processes, numerical methods for evaluating the availability of the system, and a prevision centred formulation for evaluating functionals $f \rightarrow \mathbb{E}[f(X(t))], t \in \mathbb{R}^+$, such as the availability. The space $\mathcal{X}$ is assumed to be finite.

Describing a random process corresponds to describing a random variable with an infinite dimensional co-domain, the functions. Markov chains provide a rich class of processes which can be described consistently and manipulated efficiently. We can evaluate functional $f \rightarrow \mathbb{E}[f(X(t))]$, if we know the distribution of $X(t)$.

The distribution $X(t)$ at any point in time is uniquely characterised by an initial distribution $\pi_0(x) := P(X(0) = x)$ and a family of conditional distributions $T_s^t := P(X(t) = y|X(s) = x)$. The distribution of $X(t)$ at any $t \geq 0$ may be obtained via the law of total probability as $P(X(t) = y) = \sum_{x \in \mathcal{X}} P(X(t) = y|X(0) = x)\pi_0(x)$.

A homogeneous continuous time Markov chain (HCTMC) parametrized by a transition rate matrix $Q$ (a stochastic matrix, i.e. each row sums to 0 and all non-diagonal elements are non-negative) induces a family of transition matrices $\{T_s^t : 0 \leq s \leq t\}$ such that

$$T_s^t(x, y) = \exp(Q(t-s))(x,y).$$

(4)

Together with an initial distribution $\pi_0$, Eq. (4) specifies

$$P(X(t) = y) = \sum_{x \in \mathcal{X}} \pi_0(x)\exp(Q(t))x, y).$$

(5)

Let $\mathcal{L}(\mathcal{X})$ be a space of all functions $\mathcal{X} \rightarrow \mathbb{R}$. We may also look at the transition matrix as an operator $T_s^t : \mathcal{L}(\mathcal{X}) \rightarrow \mathcal{L}(\mathcal{X})$.

For a function $f \in \mathcal{L}(\mathcal{X})$ and $x \in \mathcal{X}$, denote

$$[T_s^t f](x) := \mathbb{E}[f(X(t))|X(0) = x].$$

(6)

In the case of HCTMC, due to the homogeneity, $T_s^t f = T_s^{t-s} f$, where we will omit the subscript if it is zero, so $T_s^{t-s} := T_0^{t-s}$.

There exist multiple ways for evaluating $\mathbb{E}[f(X(t))|X(0)]$. The commonly used one is based on the additive property of probability measures. It consists of computing the
conditional probabilities $P(X(t) = y|X(0))$ for $y \in \mathcal{X}$. This allows us to construct $T'f \in \mathcal{L}(\mathcal{X})$ as $T'f(x) = \sum_{y \in \mathcal{X}} P(Y(t) = y|X(0) = x)$. The expected value of $f(X(t))$ will be $\mathbb{E}[T'f(X(0))]$. But this method is not applicable if we model the stochastic process by a non-additive measure.

Due to its definition, Eq. (4), we can decompose any transition operator into a concatenation of partial operators. For arbitrary times $t > s > r \geq 0$,

$$[T'_r f](x) = [T'_r \circ T'_s f](x). \quad (7)$$

Let $s = u_0 < u_1 < \ldots < u_n = t$ be a discretization of the time domain. Then

$$[T'_s f](x) = [T'_{u_1} \circ \ldots \circ T'_{u_{n-1}} f](x) = [T'_{u_1} f_1](x), \quad (8)$$

where $f_n = f$ and $f_i = T'_{u_{i+1}} f_{i+1}$. Therefore, we can evaluate $T'_s f$ recursively by evaluating $f_i$, starting from $f_n = f$.

We can approximate the inference by a linearized scheme up to an arbitrary precision, similarly as we do with the Euler method for solving ordinary differential equations. Given the time short is small enough,

$$\forall x \in \mathcal{X}:
\begin{align*}
&f_n(x) = f(x) \\
&\ldots \\
&f_i(x) \approx [(I + (u_{i+1} - u_i)Q) f_{i+1}](x), \quad (9) \\
&\ldots \\
&\mathbb{E}[f(X(t))]|X(0) = x] \approx [(I + (u_1 - u_0)Q) f_1](x),
\end{align*}$$

where $I$ denotes the identity operator.

V. IMPRECISE MARKOV CHAINS

Suppose that we cannot determine a single transition rate matrix $Q$ and instead have a set of ‘plausible’ matrices $\mathcal{Q}$. Set $\mathcal{Q}$ induces a set of stochastic processes $\mathbb{P}$ and we may use these as our model of the system evolution. For computing expectations of $f \in \mathcal{L}(\mathcal{X})$ over a set of stochastic processes, we need to precisely specify, which processes are present in the set of measures $\mathbb{P}$. If $\mathbb{P}$ is a set of HCTMCs with transition rates in $\mathcal{Q}$, then the lower expectation Eq. (10) is defined, in accordance with Eq. (4), as

$$\mathbb{E}[f(X(t))]|X(0) = x] = \inf_{Q \in \mathcal{Q}} \sum_{x \in \mathcal{X}} P(X(0) = x) \sum_{y \in \mathcal{X}} \exp(Qt)(x, y) f(y), \quad (10)$$

and similarly for the upper expectation. But the optimisation problem in Eq. (10) is generally intractable. Instead, we can take $\mathbb{P}$ to be a larger set of stochastic processes which comply with our knowledge. A stochastic process (not necessarily a HCTMC, see [7]) is said to be consistent with $\mathcal{Q}$ if for each $t \in \mathbb{R}_0^+$,

$$\lim_{\Delta \to 0} (P(X(t + \Delta))T_{t+\Delta} - I)/\Delta \in \mathcal{Q}. \quad (11)$$

This means that $\mathcal{Q}$ encapsulates transition rates locally, but the rates may differ over time.

Taking $\mathbb{P}$ to be a set of stochastic processes consistent with $\mathcal{Q}$, we can construct a family of sets of transition operators $\mathcal{T}'_s$ such that

$$[T'_s f](x) = \inf_{T \in \mathcal{T}'_s} \mathbb{E}[f(X(t))|X(0) = x] = \inf_{T \in \mathcal{T}'_s} \left[ I + \frac{t - s}{n} \mathbb{E}[Q f](x) \right], \quad (12)$$

where $Q : \mathcal{L}(\mathcal{X}) \to \mathcal{L}(\mathcal{X})$ is the lower transition rate operator defined for each $x \in \mathcal{X}$ as

$$[Q f](x) := \inf_{Q \in \mathcal{Q}} \mathbb{E}[Q f](x). \quad (13)$$

We can decompose the lower transition operator similarly to Eq. (7) [7] and Eq. (8) [8] in the last section, Eq. (14) allows us to construct an iterative scheme for computing lower expectations of functionals in $\mathcal{L}(\mathcal{X})$ (for more details and convergence conditions see [7] and [5]). For a discretization of the time dimension $s = u_0 < \ldots < u_n = t$,

$$f_n = f \\
\ldots \\
f_i \approx (I + (u_{i+1} - u_i)Q) f_{i+1} \quad (15) \\
\ldots \\
\mathbb{E}[f(X(t))]|X(0) = x] \approx [(I + (u_1 - u_0)Q) f_1](x),$$

VI. INFERENCE OF THE TRANSITION RATE OPERATOR

In this section, we will demonstrate how statistical inference can be used in order to construct the transition rates governing stochastic processes. We will show some results for inferences about the transition rate matrix of a HCTMC. Especially, we will focus on a Robust Bayesian inference method proposed in [8].

Suppose that we observe the trajectory, say $\omega$, of a Markov chain. Denote $d_x$ the total holding time ($\int_0^t X(t) dt$) in state $x$ and $n_{xy}$ the number of transitions from state $x$ to state $y$. For the inference about the transition rate matrix generating the stochastic process, we may use the following likelihood results [6]

$$\mathcal{L}(\omega|Q) = \prod_{x \neq y}^{\text{state}} (q_{xy})^{n_{xy}} \exp(-q_{xy}d_x), \quad (16)$$
which leads to the Maximum Likelihood Estimator $q_{xy}^{\text{MLE}} = \frac{n_{xy}}{n_x}$. 

Because the likelihood belongs to the exponential family, it was proposed in [8] to use a product of gamma priors for the Bayesian inference. For a chosen collection of hyperparameters $\alpha = \{\alpha_{xy}, x \neq y \in X\}$ and $\beta = \{\beta_x, x \in X\}$, the prior density $\pi$ is

$$\pi(Q|\alpha, \beta) = \prod_{x,y \in X \overline{x \neq y}} (q_{xy})^{\alpha_{xy}-1} \exp(-q_{xy} \beta_x)$$

$$= \prod_{x,y \in X \overline{x \neq y}} \Gamma(\alpha_{xy}) q_{xy}^{\alpha_{xy}-1} \beta_x. \quad (17)$$

Combining the prior, Eq. (17), with the likelihood, Eq. (16), we can pose the exact form for the posterior mean, $E[q_{xy}|\alpha, \beta] = \frac{\alpha_{xy} + n_{xy}}{\beta_x + d_x}.$ \quad (18)

To avoid the necessity to choose any particular prior distribution, we may, instead, select a whole set of them. This method is known as a Generalized Bayesian inference [12] and will generally lead to a set of posterior distribution, hence a set of posterior mean estimates and an Imprecise Markov Chain.

In [8], a set of prior distributions based on the imprecise Dirichlet model (IDM) (also [12]) has been proposed. Among the advantages of the IDM model is the property of prior ignorance. Before observing any samples, the IDM generates vacuous inferences (all events have lower probability 0 and upper probability 1 - this is impossible to achieve with precise priors).

In combination with Eq. (16) and Eq. (18), this model leads to a tractable lower transition rate operator, Eq. (13). For an arbitrary ancillary random variable $L$, we can decompose the expected value in Eq. (2) via the law of total expectation into

$$E[\varphi(X)] = E\{E[\varphi(X)|L]\}. \quad (21)$$

Suppose that we can group the components of the system such that states of the components of the same group are exchangeable. We may choose the ancillary random variable $L$ to be a vector of numbers of functioning components of each group (arbitrarily ordered) and define the system survival signature as $\Phi(l) = E[\varphi(X)|L = l]$. Due to the exchangeability assumption, the conditional probability $P(X|L)$ is uniform on the subset of $X$ for which $L = l$. This allows us to compute $\Phi$ according to the laws of classical probability as the fraction of functional states,

$$\Phi(l) = \frac{|\{x \in X : L(x) = l, \varphi(x) = 1\}|}{|\{x \in X : L(x) = l\}|}, \quad (22)$$

where the notation $L(x)$ represents the dependency of $L$ on the actual component states.

In this section we will present some methods which help us to decrease the computational complexity arising from the combinatorial explosion of the system components’ state space. Since each component is binary and arbitrary combination of component states is permitted, the cardinality of the state-space $X$ would be $2^N$, where $N$ is the number of components in the systems. This itself often makes it intractable to assess reliability for some real systems even without taking into account the dependencies influencing their deterioration or imprecise probability models.

A. Survival Signatures

Survival signatures were introduced to facilitate analyses of large heterogeneous systems [3]. They are used to simplify the prescription of the relation between the system state and the component states. Assuming some simplifying relations among the components interactions, exchangeability of their states, we can compress the remaining relevant aspects of the structure function into a lower dimensional summary. If the exchangeability assumption holds, we can still recover exact inference about the system reliability.

These results require that every state had been observed, i.e. that $\forall x \in X : d_x > 0$. This may be overly restrictive requirement for application of the method in realistic scenarios from reliability theory and will be resolved in Section VIII.
B. State-Space Lumping

Lumping serves as a way to decrease sizes of state-spaces of Markov models by grouping some of the states together. But, if the aim is to obtain a Markov Chain again after the lumping, its uses are limited and the existence of a limped HCTMC is not always guaranteed \[2\]. In general, the lumped process may not remain Markov or homogeneous after the procedure. Generally, it will become a Homogeneous (generated by a single set of transition rate operators \(\mathcal{D}\) at each time) Imprecise Markov Chain \[4\].

Let \(Q\) be a transition rate matrix of a (precise) HCTMC with state space \(\mathcal{X}\). Let \(\Lambda : \mathcal{X} \to \hat{\mathcal{X}}\) be a surjective function mapping states from the original state-space on their respective groups in the new state-space. The evolution of the lumped state \(\hat{X}(t) = \Lambda(X(t))\) will be induced by the evolution of state \(X\) according to the original HCTMC.

Let \(f \in \mathcal{L}(\mathcal{X})\) be lumpable w.r.t. \(\Lambda\), meaning that there exists \(\hat{f} \in \mathcal{L}(\hat{\mathcal{X}})\) s.t.

\[
\forall x \in \mathcal{X} : f(x) = \hat{f}(\Lambda(x)).
\]

The transition rate operator \(\hat{Q} : \mathcal{L}(\hat{\mathcal{X}}) \to \mathcal{L}(\hat{\mathcal{X}})\) of the lumped process can be evaluated as

\[
(\hat{Q} \hat{f})(\hat{x}) = \min_{x \in \Lambda^{-1}(\hat{x})} \left\{ \sum_{y \in \mathcal{X}} \hat{f}(\hat{y}) \sum_{y \in \Gamma(y)} Q(x, y) \right\}. \tag{24}
\]

C. Combining Lumping and Survival Signatures

Since lumping and signatures share similar principles and goals, they can be combined for the purposes of Markov analysis. Lumping will allow us to simplify the joint state-space of the system components and survival signature will provide a method for constructing a lumpable function which we will use for assessing the system reliability.

In order to utilize the survival signatures, we need to assure exchangeability of the states of the grouped components. Clearly, for any Markov model of the component degradation, we may consider each component of its own distinct type. But such a model would lead to no savings of computational resources. Since we intend to perform statistical inference based on observations of independent copies of a system, we will include assumptions in our statistical model to make the inference possible and tractable.

Let us aggregate the system components into several groups. A simple assumption which will make the combination possible is to assume that the transition rates depend only on the number on functioning components of each group. An example of such Markov model is depicted on Figure 1. The example also demonstrates, that this assumption leads to a precise HCTMC after lumping, which we currently need in order to carry out statistical inference as described in Section \[VI\].

This assumption also ensures exchangeability of the component states. If we therefore use \(\Lambda(X) = L(X)\), as used in Section \[VII-A\] as the lumping map, and \(\Phi \circ L \in \mathcal{L}(\mathcal{X})\) as the lumping function, we may use the procedures in Eq. \[9\] and Eq. \[15\] on the lumped process to assess the system reliability (= \(\mathbb{E}[\Phi(L(X))]\)).

VIII. FORCED SEMI-VACUOUS MODEL WITH STRUCTURAL ASSUMPTIONS

The drawback of the methods mentioned in Section \[VI\] the MLE and Eq. \[20\], is the necessity to observe all the possible states \(x \in \mathcal{X}\). Otherwise the null value of the denominator \(d_x\) will cause that the expressions will not be properly defined. Given the cardinality of the state-space in reliability problems and often a low number of observations, it is highly unlikely that all the states will be observed. Although we have demonstrated some methods to reduce the size of the state-space in the Section \[VII\] it still remains unlikely to observe all the states even from this smaller state-space in practical applications.

Recall the linearized computation scheme for assessing lower expectations of functionals of stochastic processes, Eq. \[15\]. If we replace the linear approximation by the original expression it approximates, the general step of the procedure may be viewed as

\[
f_i(x) = [\sum_{u_i} f_{i+1}](x). \tag{25}
\]

While evaluating \(f_i\) for each of the states \(x \in \mathcal{X}\) during the step, two situations may occur based on the data from which we draw our inferences about the underlying transition rate matrix. Either state \(x\) was observed, in which case we may use the original scheme with \(Q\) given by Eq. \[20\]. If state \(x\) was not observed, then we know nothing about transitions from \(x\), thus nothing about \(\mathbb{E}[X(t) = y | X(s) = x]\) for any \(y, s, t\). But we can reflect this ignorance in the IP models simply by calling for the properties of the transition operator for which \(\sum f \geq \inf f\).

But a bare inclusion of vacuous model may lead to excessively imprecise inferences. There is still an additional information

\[ L0, L1, L2, L3 \]

\[ 000, 001, 010, 011 \]

\[ 100, 101, 110, 111 \]

\[ X0, X1, X2, X3 \]

\[ \hat{X0}, \hat{X1}, \hat{X2}, \hat{X3} \]

\[ \hat{Q} \]

\[ \hat{f} \]

\[ \hat{\mathcal{X}} \]

\[ \Lambda \]

\[ \mathcal{L}(\mathcal{X}) \]

\[ \mathcal{L}(\hat{\mathcal{X}}) \]

\[ \mathbb{E}[\Phi(L(X))] \]

\[ \sum \]

\[ \inf \]

\[ \geq \]

\[ \hat{X}(t) = \Lambda(X(t)) \]

\[ \Lambda^{-1}(\hat{x}) \]

\[ \hat{f}(\hat{y}) \]

\[ \sum_{y \in \Gamma(y)} Q(x, y) \]

\[ \Lambda(X) = L(X) \]

\[ f_i(x)\]

\[ \sum_{u_i} f_{i+1}(x) \]

\[ \mathcal{D} \]

\[ \mathcal{X} \]

\[ \mathcal{L}(\mathcal{X}) \]

\[ \mathcal{L}(\hat{\mathcal{X}}) \]

\[ \mathbb{E}[\Phi(L(X))] \]

\[ \sum \]

\[ \inf \]

\[ \geq \]

\[ \hat{X}(t) = \Lambda(X(t)) \]

\[ \Lambda^{-1}(\hat{x}) \]

\[ \hat{f}(\hat{y}) \]

\[ \sum_{y \in \Gamma(y)} Q(x, y) \]

\[ \Lambda(X) = L(X) \]

\[ f_i(x)\]

\[ \sum_{u_i} f_{i+1}(x) \]

\[ \mathcal{D} \]

\[ \mathcal{X} \]

\[ \mathcal{L}(\mathcal{X}) \]

\[ \mathcal{L}(\hat{\mathcal{X}}) \]

\[ \mathbb{E}[\Phi(L(X))] \]

\[ \sum \]

\[ \inf \]

\[ \geq \]
we can supply to the model in order to improve our inferences. That is by introducing additional structural assumptions. It is not too preposterous, in many reliability applications, to assume that no more than one failure may occur at any time instance. Exceptions could of course be found if we were to include possibilities of e.g. common cause failures. Nevertheless, even in those cases, the method we are going to introduce might still be applied at least partially.

Assume that for each of the states $x \in \mathcal{X}$, we can rule out some states $W(x) \subset \mathcal{X}$ for which we deem the transition $x \rightarrow y \in W(x)$ to be impossible. In the case of non-increasing component state processes, this would be e.g. transitions that would renew some component’s functionality. If we judge that only one failure can happen at any time instant, then we could rule out all the transitions which would render two or more components non-functional at the same time. In any case, such an assumption will have an effect on both the lower transition rate operator emerging from the robust Bayesian inference (Eq. (20)) and also on the vacuous prevision.

In the selection of the set of prior distributions for the robust inference, leading to the set of posterior rate matrices in Eq. (19), we can narrow down the set of admissible transition operators $\mathcal{T}$. If we rule out transitions $W(x)$ for state $x$, we can reflect this in the set of admissible transition operators by selecting $\tilde{\mathcal{T}}(x) := \{A(x,y) \in \mathcal{T} : A(x,z) = 0, \forall z \in W(x)\}$. This would lead to an enhanced posterior transition rate operator

$$
\tilde{\mathcal{Q}}(f)(x) := \sum_{x \in \mathcal{X}} \min_{y \in \mathcal{X} \setminus W(x)} (f(y) - f(x)) + \sum_{y \in \mathcal{X} \setminus \{x\} \cup W(x)} \frac{n_{xy}}{d_x} (f(y) - f(x)).
$$

(26)

The introduced assumption will also affect the vacuous prevision $T \geq \inf f$ simply by taking the infimum only over the set of admissible transitions. Hence $\tilde{T}(f)(x) \geq \inf_{y \in \mathcal{X} \setminus W(x)} f(y)$. Therefore, we redefine the step in the linear approximation scheme in Eq. (15) as

$$
\tilde{\mathcal{T}}^{u_{i+1}} f(x) \approx \begin{cases} 
[f + (u_{i+1} - u_i) \tilde{Q}] f(x), & d_x > 0, \\
\min_{y \in \mathcal{X} \setminus W(x)} f(y), & d_x = 0.
\end{cases}
$$

(27)

Now, the inference is always possible.

**IX. Example**

We will introduce the methodology on a simple example. We assume the system topology as depicted on Figure 2. Components of type 1 are represented by the white blocks and components of type 2 are the green ones.

HCTMC, including additional assumptions that 1) component state processes are non-increasing, and 2) only one failure may occur at any time instant, is depicted on Figure 3. The arrows represent the only admissible transitions.

We simulate several observations of the system lifetime. The simulated experimental procedure produce evolution of the system components’ states from which we can extract $d_x$ and $n_{xy}$ needed for the likelihood function, Eq. (16). The simulation is terminated by the time the system reaches a failure state, therefore we never observe all the states in the state-space.

The inference of the survival function (Eq. (2)) is computed via the linearized scheme (Eq. (15)) as $	ext{Rel}(t) = \mathbb{E}\left[\Phi(L(t))|L(0) = \mathbf{M}\right] = \left[\tilde{T}^{\Phi}\right](\mathbf{M})$. $\mathbf{M}$ represents vector of number of components of each type in the system, hence $L(0) = \mathbf{M}$ implies that all the components are functional at time 0, according to an assumption we have made in Section III. Analogously it is done for the upper bound.

Resulting inferences of the bounds on the survival function are show on Figure 4. The inference seems to properly envelope the ground truth (GT) survival function, exhibit great imprecision if the data is scarce (the first plot with 5 observations), and approaches the ground truth survival function as the number of observations increases.

**X. Conclusion**

Imprecise Markov Chains present a natural way for conducting Markov analysis in reliability theory. They allow us to robustify our inferences by including multiple plausible models of the systems’ dynamics. So far, it seems that they provide the only way to conduct statistical inference in this field because they enable to include vacuous models to substitute for the lack of observations. Imprecise Markov Chains also provide a general lumping framework for dimension reduction for large state-spaces. Combined with the Survival Signature framework, they present a powerful analytic technique for conducting inferences on large - real world systems. Although
Fig. 3. Lumped transition diagram of the example system. Only the depicted arrows signify admissible state transitions of the system.

<table>
<thead>
<tr>
<th>$t_1$</th>
<th>$t_2$</th>
<th>$\Phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0.4</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.033</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.166</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0.466</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>0.066</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0.1066</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0.2</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>0.3866</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>0.66</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>0.2</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>0.31</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>0.47</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0.67</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>0.85</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>4</td>
<td>0.4</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>0.56</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>0.74</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>0.88</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>0.9466</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>5</td>
<td>0.66</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>0.8</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>0.933</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>1</td>
</tr>
</tbody>
</table>

TABLE I
SURVIVAL SIGNATURE OF THE EXAMPLE SYSTEM. NULL CONTRIBUTIONS ARE OMITTED.

Fig. 4. Inferred bounds for the survival function of the example system for $n=5, 15, 50$, and 150 observations (depicted respectively from top to bottom).
the method was demonstrated only on a small system, both lumping and survival signatures are being used to enable analyses which would be intractable otherwise due to their high dimensionality. Besides that, Imprecise Probability models allow us to conduct Robust Bayesian inference, which mitigate some of the controversies about selecting prior distributions in Bayesian inference.

An interested reader is encouraged to follow literature references provided in relevant sections for a more complete and rigorous treatment of the addressed topics.

ACKNOWLEDGMENT

This work is funded by the European Commissions H2020 programme, through the UTOPIAE Marie Curie Innovative Training Network, H2020-MSCA-ITN-2016, Grant Agreement number 722734.

REFERENCES