In Supremae Dignitatis · Università di Pisa

FACOLTÀ DI INGEGNERIA
Dipartimento Di Sistemi Elettrici ed Automazione

Dependability Analysis of a Safety Critical System: The LHC Beam Dumping System at CERN

Tesi di Dottorato di Ricerca in Automatica, Robotica e Bioingegneria

Relatore:
Prof. Ing.
ALDO BALESTRINO

Candidato:
Ing.
ROBERTO FILIPPINI

Co-relatore:
Dr. Ir.
JAN UYTHOVEN

Anno Accademico 2005-2006

Abstract

This thesis presents the dependability study of the Beam Dumping System of the Large Hadron Collider (LHC), the high energy particle accelerator to be commissioned at CERN in summer 2007. There are two identical, independent LHC Beam Dumping Systems (LBDS), one per LHC beam, each consisting of a series of magnets that extract the particle beam from the LHC ring into the extraction line leading to the absorbing block. The consequences of a failure within the LBDS can be very severe. This risk is reduced by applying redundancy to the design of the most critical components and on-line surveillance that, in case of a detected failure, issues a safe operation abort, called false beam dump.

The system has been studied applying Failure Modes Effects and Criticality Analysis (FMECA) and reliability prediction. The system failure processes have been represented with a state transition diagram, governed by a Markov regenerative stochastic process, and analysed for different operational scenarios for one year of operation. The analysis of the system results in a safety level ranked SIL4 in the IEC 61508 standard and 4 (± 2) expected false beam dumps generated per LBDS. These results will be validated through a three months reliability run. Several sensitivity analyses have been made providing additional evidence on the importance of the fault tolerant design features and the achieved trade-off between safety and availability.

The Beam Dumping System is part of the LHC machine Protection System for which a safety level SIL3 is required. A simplified model of the LHC Machine Protection System (MPS), including the LBDS and other critical protection systems, has been analysed. Depending on the hazards (e.g. the fast beam losses being the most critical event in the LHC) and their coverage, the safety of the MPS has been calculated between SIL2 and SIL4 with about 40 (± 6) expected false dumps per year, which is the 10% of the machine fills. In the context of the MPS the LBDS is one of the safest systems and contributes to unavailability with an acceptable fraction of false dumps.

Acknowledgements

This thesis is the fruit of three years of work conducted at CERN in the AB/BT group. My first sincere thanks go to my supervisor Dr. Jan Uythoven. He has provided to me a continuous support during my stay at CERN, contributing in a significant way to the progress of my research with suggestions and ideas that have improved the thesis in its present version. I wish to thank Prof. Ing. Aldo Balestrino who has been my tutor at the Pisa University, for his expert guidance and the positive vision he has always had on my work. I owe gratitude to my colleagues of the AB/BT group at CERN, especially the group leader Dr. Volker Mertens and the project leader of the LHC Beam Dumping System, Dr. Brennan Goddard. I am pleased to thank Dr. Rudiger Schmidt for the precious opportunity that he gave me to apply my dependability studies also on the larger context of the LHC machine protection.

Among the other people, I wish to thank a colleague and friend of mine, Dr. Angelo Alessandri who has contributed to the revision of part of my thesis since the very early stages. A thanks goes also to my summer student David Huwe Jones, for the good work he performed during his short stay at CERN.

Finally, I wish to thank Prof. Fabio Uccelli, former employee of the INFN, who was at CERN in the sixty and Prof. Lorenzo Foa of the Scuola Normale of Pisa who encouraged me to have the unique experience of working at CERN.

Contents

List of Figures

List of Tables

Introduction

The Large Hadron Collider (LHC), approved by the CERN Council in December 1994, is planned to come into operation by the end of 2007. The accelerator occupies an approximately circular tunnel, 27 km in circumference and 3.8 m in diameter, located between 100 m and 150 m underground, crossing the Swiss-French border at the periphery of Geneva. The LHC will accelerate two counter-rotating beams of protons that are nuclei of hydrogen atoms [56, 57]. The two beams will collide at a centre of mass energy of 2×7 TeV, about 35 times the energy of the LEP [55], a previous accelerator at CERN, and 7 times the energy of the Fermilab Tevatron [64], which makes the LHC the world's most powerful particle accelerator for high energies physics. The collisions will result in the scattering and disintegration of the nuclei and their constituents with the production of particles that will permit to investigate the matter on a sub-nuclear scale, searching for signatures of super-symmetry, dark matter and the origins of mass [56].

This thesis presents the analysis of the dependability of the LHC Beam Dumping System (LBDS). The LBDS performs the extraction of the high energy proton beam from the LHC ring, its dilution and steering through the dump tunnel and the safe deposition into an absorber block. Any failure leading to the unavailability of the LBDS during the operation is a severe safety concern for the LHC. The system has been designed in order to contain the residual risk of failure. A safety level of SIL3 of the IEC 61508 standard is demanded for the LHC Machine Protection System, of which the LBDS is an important component. The study follows some similar work on different equipment already done at CERN (e.g. the quench protection system [90] and the beam loss monitor system [37]). These studies describe the consequences of failures and give an estimate of their likelihood, which is the ultimate goal of the dependability assessment. They all apply FMECA (Failure Modes, Effects and Criticality Analysis) though they differ in the methodology for the analysis: the Monte Carlo simulation was applied in [90] while fault tree analysis was used in [37].

This study applies Markov processes and Markov regenerative stochastic processes to the modeling and analysis of dependability problems. This approach is demonstrated to be an elegant and mathematically exact way to describe the system failure processes together with the dependability attributes, as alternative to either a fault-tree or a Monte Carlo approach.

The work is organized as follows. Chapter 1 gives an overview of the LHC accelerator and some rudimentary information on accelerator physics and technology. The LBDS is described in Chapter 2. Chapter 3 introduces dependability terminology and design methods that apply to make a system resilient to failure and safe in particular. Chapter 4 describes the probability models for modeling the dependability, which are used in Chapter 5 for reliability, availability and safety applications. Chapter 6 outlines the FMECA analysis of the LBDS. In Chapter 7 the model of the LBDS failure processes is built and an analysis is performed for one year of operation and different operational scenarios, resulting in figures for safety and availability, which are completed by an additional sensitivity analysis of the main design parameters. Chapter 8 provides an overall estimate of the dependability for a simplified LHC machine protection system, including the LBDS and the most important protection systems. A summary of the results of the study and some final remarks are given in Chapter 9.

Chapter 1

LHC Overview

1.1 The LHC Accelerator

The LHC is the highest energy accelerator of a chain of accelerators, ranging from the particle production to the injection in the LHC rings where the beam is stored, accelerated and finally extracted at the end of the operational cycle, see Figure 1.1. In each accelerator the beam energy is increased [56]. The protons p+ are a type of hadrons, which form a broad category of particles that includes also neutrons and in general all particles that build the nucleus of the atoms[1]. They are produced in the LINAC (LINear ACcelerator), packed in bunches during the acceleration process and transferred to the PS booster and further into the PS (Proton Synchrotron). From the PS they are transferred to the SPS accelerator (the Super Proton Synchrotron) where they are accelerated up to an energy of 450 GeV and subsequently injected into the LHC rings[2]. The final LHC proton beam, at nominal intensity, will consist of 2808 (n_b) bunches , each containing 1.15×10^{11} protons (N_p) resulting in a total beam current of 0.584 A. The two beams are accelerated to a beam energy of 7 TeV and can be kept circulating for hours at the ultra-relativistic velocity of 0.999999991 times the speed of light, completing

[1] The LHC will also accelerate ions but at less intensity than the proton beams.

[2] The nominal LHC filling requires 12 injections from the SPS for each LHC ring [57].

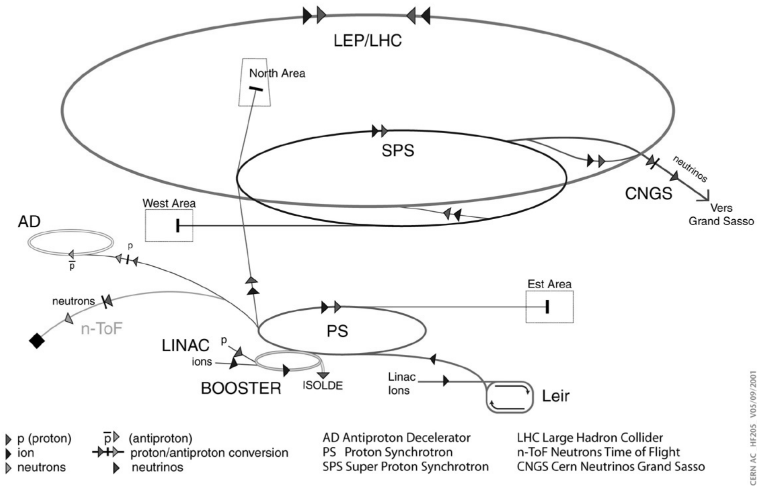

LEP/LHC
North Area
SPS
neutrinos
CNGS
Vers
Grand Sasso
West Area
AD
p
p̄
neutrons
n-ToF
Est Area
PS
LINAC
p
ions
BOOSTER
ISOLDE
Linac
Ions
Leir
p (proton)
ion
neutrons
p̄ (antiproton)
proton/antiproton conversion
neutrinos
AD Antiproton Decelerator
PS Proton Synchrotron
SPS Super Proton Synchrotron
LHC Large Hadron Collider
n-ToF Neutrons Time of Flight
CNGS Cern Neutrinos Grand Sasso
CERN AC HF205 V05/09/2001

Figure 1.1: The CERN complex [19].

a machine circumference of 26.7 km in 89 μs.

Beam collisions are foreseen at four interaction points in the heart of the main experiments: ALICE (A Large Ion Collider Experiment), ATLAS, CMS (Compact Muon Solenoid), and LHCb (LHC beauty experiment). At the ATLAS and CMS experiments, the beams are squeezed in transverse beam size to about 16 μm, which increases the chances of a collision among the individual particles of the two beams. Just to give an idea, the squeezing of 100,000 million protons (at nominal beam currents) per bunch down to 16 μm (1/5 the width of a human hair) at an interaction point results in around 20 collisions per crossing. The bunch spacing is 25 ns, so that one collision occurs every 25 ns. The LHC main parameters are listed in Table 1.1.

1.2 Accelerator Physics and Technology

The LHC is a synchrotron, which is a type of particle accelerator usually characterized by a quasi-circular vacuum chamber (i.e. the ring) in which the beam is circulating, surrounded by magnets. The ring is split in 8 octants, separately powered and accessible from the surface through the access points, see Figure 1.2. One beam is injected at point 2, the other at point

Quantity	**Value**
Beam energy (E)	7.00 TeV
Beam current	0.584 A
Circumference	26.7 Km
Number of protons per bunch (N_p)	1.15×10^{11}
Dipole field (B)	8.4 Tesla
Revolution frequency (f_{rev})	11.24 kHz
Number of bunches per beam (n_b)	2808
Proton (rest) mass (m_0)	1.672×10^{-27} kg
Normalized emittance (ϵ_n)	3.75 μm × rad
Proton charge (q)	1.602×10^{-19} Coulomb
Beta function at collision (β)	0.50 m
Luminosity (L)	1.0×10^{34} $cm^{-2}sec^{-1}$

Table 1.1: The LHC and beam main nominal parameters [56].

8. The extraction or beam dumping system is located at point 6. The Radio Frequency (RF) system, required to accelerate the beam, is placed at point 4 while the collimators (beam cleaning) are at points 3 and 7.

The LHC ring is occupied by magnets for steering and focusing the beam in order to keep a high energy and intensity beam circulating for the necessary length of the experiments. The most important magnets for the LHC are the dipole magnets and the quadrupole magnets.

The **dipole** magnets play a crucial role in the LHC as they determine the bending angle of the quasi circular orbit in the horizontal plane. The relation between bending angle and magnetic field of the dipole magnets is calculated with **Lorenz's law**. The bending angle Θ is determined by the magnetic induction B, the relativistic mass m, the charge q, the velocity v of the particle and the length of the magnet l:

$$\Theta = q \times \frac{Bl}{mv} \tag{1.1}$$

where $m = \frac{m_0}{\sqrt{1-(v/c)^2}}$. At the top LHC beam energy of 7 TeV the required magnetic induction is 8.4 T for 14.3 m long magnets, resulting in a bending angle of 0.0051 rad. This field requires a current of around 11,700 A in the

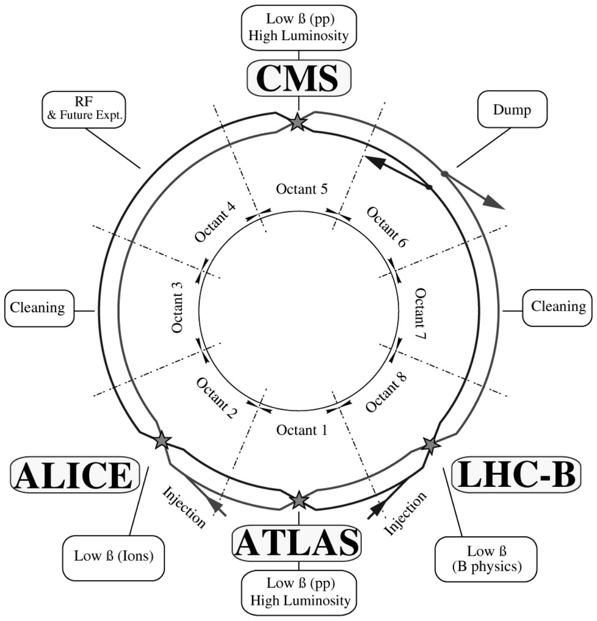


Figure 1.2: LHC schematic.

superconducting dipole magnets for a total of 1232 of magnets that deviate the beam over 2π. The super-conducting magnets of the LHC work at a super-fluid helium temperature of 1.9 K. They have two apertures, one for each of the counter-rotating beams. In order to avoid undesired collisions with residual gas molecules, an ultra high vacuum of 1.33×10^{-10} mbar ($\cong$ 3 million molecules/cm^3) is created in the beam pipes. Additional small orbit correctors, shorts dipole magnets, are installed to correct the beam orbit in the horizontal and vertical planes.

The **quadrupole** magnets focus the beams in the transverse planes. A quadrupole magnet has four poles with alternating polarities, symmetrically arranged around the centre of the magnet. The resulting magnetic field lines follow a hyperbolic contour with the strength increasing proportionally to the distance from the centre. The magnet acts like a focusing lens in one plane, and a defocusing lens in the other plane. A global beam focusing in both planes is reached by alternating focusing and defocusing quadrupole magnets, called FODO lattice. The obtained transverse motion around the circular trajectory for a single particle is described by **Hill's equation**.

$$\frac{d^2z}{s^2} + K(s)z = 0 \tag{1.2}$$

where z stands for either the horizontal x or vertical coordinate y and s is the longitudinal displacement along the reference orbit[3]. The focusing strength $K(s)$ is a function of s, $K(s) = (q \times g(s))/p$, where $g(s) = \partial B_y/\partial x$ is the gradient of the quadrupole magnet in the x plane and $g(s) = -\partial B_x/\partial y$ is the gradient in the y plane, p is the momentum and q is the charge of the particle. The solution of (1.2) is a residual harmonic oscillation called betatron oscillation:

$$x(s) = \sqrt{\epsilon\beta(s)}cos(\phi(s) + \phi_0) \tag{1.3}$$

where ϵ is the a constant called emittance[4], which describes the beam quality, $\beta(s)$ is the beta-function, $\phi(s) + \phi_0$ is the phase advance which together describe the magnetic optics. The equation (1.2) is solved for the initial condition ϵ and ϕ_0, which define the position of the particle in the phase space x, dx/ds. Other quantities derived from (1.2) are the number of betatron oscillations per turn that is called the betatron tune Q, and the transverse beam size which is given by $\sigma_{x,y} = \sqrt{\epsilon\beta}$.

The beam is accelerated by a longitudinal electric field from a resonant cavity, generated by the Radio Frequency (RF) system. The frequency of the cavity is set to a multiple of the beam revolution frequency. In the LHC, this frequency is 400 MHz, which corresponds to an oscillation period of 2.5 ns. This defines the stable phase for the bunches of particles, separated by 25 ns at the injection in the LHC and during the complete LHC operational cycle. All particles describe a longitudinal oscillation around the stable phase, called synchrotron oscillation (for the LHC around 21.4 Hz). The synchrotron

[3] This equation is only valid for particles without any energy deviation and large bending radii.

[4] The emittance is related to the normalized emittance as quoted in Table 1.1 by $\epsilon_n = \beta\gamma\epsilon$, where β and γ are the relativistic parameters.

oscillation must not be confused with another phenomenon, the synchrotron radiation, which is related to the deflection of the high energy particles by the dipole bending magnets, giving rise to emission of light. The synchrotron radiation is mostly important for light particles like electrons and was a major effect for the LEP. Nevertheless, it will not be totally negligible for the LHC due to the high energy reached for the beams. The RF will compensate these losses.

Second order phenomena exist that affect the beam quality and require the addition of higher order multipole magnets. Sextupole magnets are placed close to the quadrupoles in the LHC in order to control the chromaticity, a quantity that relates the Q tune spread to the momentum spread of the beam. The skew quadrupoles control the coupling between the two transverse planes. Other magnets like octupoles and decapoles are introduced in order to handle resonances and higher order effects. A more detailed description of this extensive subject is beyond the scope of this thesis and is not treated here.

All magnets for steering and focusing the beam or compensating the higher order phenomena are arranged in a cell structure that repeats identically along the arcs of the LHC ring, see Figure 1.3. Every three bending dipoles there is a quadrupole with orbit correctors positioned close to it.

Special types of magnets are used for the injection and the extraction of the two beams into and from the LHC ring respectively. These are the kicker magnets, which are pulsed, and the septum magnets, which work continuously. The extraction magnets are part of the beam dumping system. In this system there are 15 extraction fast pulsed kicker magnets, followed by 15 septum magnets and 10 dilution kicker magnets that guide the beam along 700 meters of the dump tunnel toward the graphite absorber block. The safety of the beam dumping system of the LHC is the main subject of this thesis. Details of its functioning and the design will be given in Chapter 2.

The most important LHC parameter for the experiments is the **beam**

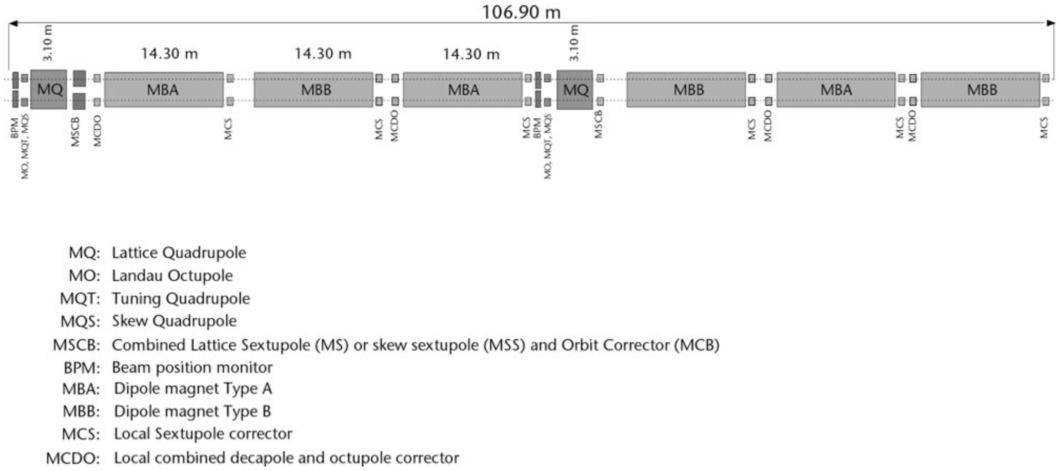

106.90 m
3.10 m
14.30 m
14.30 m
14.30 m
3.10 m
MQ
MBA
MBB
MBA
MQ
MBB
MBA
MBB
BPM
MO, MQT, MQS
MSCB
MCDO
MCS
MCS
MCDO
MCS
BPM
MO, MQT, MQS
MSCB
MCS
MCDO
MCS
MCDO
MCS
MQ: Lattice Quadrupole
MO: Landau Octupole
MQT: Tuning Quadrupole
MQS: Skew Quadrupole
MSCB: Combined Lattice Sextupole (MS) or skew sextupole (MSS) and Orbit Corrector (MCB)
BPM: Beam position monitor
MBA: Dipole magnet Type A
MBB: Dipole magnet Type B
MCS: Local Sextupole corrector
MCDO: Local combined decapole and octupole corrector

Figure 1.3: Regular arc cell of the LHC [19].

luminosity. This is defined as the number of particles per square centimeter, per second crossing at the interaction point:

$$L = \frac{N_p^2 n_b f_{rev}}{2\pi\sigma^2} \tag{1.4}$$

Whereas in past and present colliders the luminosity culminates around $L = 10^{32} cm^{-2} s^{-1}$, the LHC is designed to reach $L = 10^{34} cm^{-2} s^{-1}$, two orders of magnitude higher, see also Table 1.1. This gain in luminosity is not for free. When two bunches cross in the centre of a physics detector only a tiny fraction of the particles collide, giving about 20 collisions per crossing, and result in the wanted events. The large majority of the particles are deflected by the electromagnetic field of the opposing bunch without colliding. These deflections, which are stronger for denser bunches, accumulate turn after turn and may eventually lead to particles loss. This beam-beam effect was studied in previous colliders, where experience showed that one could not increase the bunch density beyond a certain value, the so called beam-beam limit, to preserve a sufficiently long beam lifetime. In order to reach the maximum luminosity the LHC has to operate as close as possible to this limit.

A collimation system, made of various blocks of different materials, is designed to catch and absorb any unstable particles before they can reach the beam pipe wall and quench[5] the superconducting magnets. This way the beam losses are confined in well-shielded regions far from any superconducting elements. Beam losses in the LHC ring are detected by the beam loss monitoring system or indirectly by the quench protection system of the superconducting magnets. Other instruments are also required, in particular the Beam Position Monitors (BPM) to measure the beam position.

1.3 Safety Concerns and Machine Protection

The studies on the safety of particle accelerators are a necessity for the accelerators with large stored beam energy for which the severity of failures are drastically amplified. In this respect, the LHC is going to be the most powerful particle accelerators both for the energy stored in the beam and the energy stored in its magnets[6]. The consequences of a failure in the LHC are estimated in damage to costly superconducting magnets and radioactive contamination, resulting in many months or even years of downtime for the accelerator. For these reasons, the LHC project is demanded to comply with strict safety recommendations similar to those applied in nuclear power engineering[7].

The safety in the LHC is assured by the LHC Machine Protection System MPS [56, 92]. The MPS checks continuously for the existence of safe conditions for the LHC, before entering operation and especially during operation when the beam is circulating. In case of detected failures in the machine or beam anomalies, the MPS issues a beam dump request and the operation is aborted. The unavailability of the MPS has potentially serious consequences [89], depending on the part of system which failed or is left without protec-

[5] The quench is the transition of the superconducting state to its normal state and releases the magnetic energy stored in the magnets.

[6] The energy stored in the magnets is 11 GJ [56], equivalent to 2.8 tons of TNT.

[7] The timescale of failures developing in a nuclear power plant is longer than that in the particle accelerators where a reaction time of milliseconds could not be sufficient to preserve the machine from the catastrophe.

tion and exposed to an increased hazard. The overall MPS is required to be SIL3, that corresponds to a failure rate in the interval $[10^{-8}/\text{h}, 10^{-7}/\text{h}]$ as specified in the IEC 61508 standard [42]. The failsafe strategy of aborting the operation in case of any detected failure is expected to be determinant in the achievement of the required safety level for the LHC, as it will be explained in section 3.5. Nevertheless, none of the MPS components should disrupt the machine operation above a reasonable limit by creating physically unfounded dump requests or beam-inhibit signals. This is a further requirement which determines a trade-off between the safety and availability for the MPS.

The MPS consists of a large number of complex systems involved in the protection task of the LHC. Some of these systems are devoted to the beam surveillance (e.g. beam losses, beam position, etc.), others to the surveillance of the status of critical equipment (e.g. super conducting magnets, power converters, etc.). An inventory of these includes: the beam loss monitor system to detect beam losses, the beam dumping system for the beam extraction, the quench protection system that detects load change in the superconducting elements, the powering interlock controller that interlocks the powering of the superconducting magnets, the collimation systems, the RF system, the beam position monitors, the vacuum system and others (see Figure 1.4).

The core of the MPS is the Beam Interlocking System that consists of 16 Beam Interlock Controllers (two per sector) communicating via fiber optics to form the beam permit loop, see Figure 1.5. All machine protections systems are directly or indirectly connected to a Beam Interlock Controller (BIC) with their user permit signal. The beam permit loop is the result of a handshaking protocol (user permit $\Longleftrightarrow$ beam permit) between the users and the local BIC. Each BIC receives the 10 MHz signal (i.e. the token) that is retransmitted to the neighbour BIC only if all local user's permit signals are received, which means that they are functioning. If this holds for all 16 BICs, the token starts circulating and the beam can be injected in the machine. For reliability reasons, there are two loops per beam, four in total, with one token

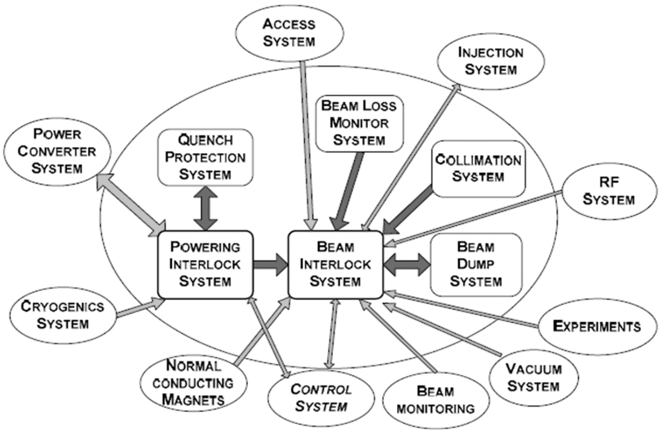


Figure 1.4: The Machine Protection Systems along the LHC Ring [56].

circulating in the clockwise direction and another anti-clockwise. Each BIC may cut the loops at any moment during operation in case at least one of the user permits has turned to false[8], and a dump request is transmitted to the LBDS. The timing response from the detected critical event to the transmission at the LBDS interface is estimated to be about 70 μs [13].

Among the systems that make part of the MPS, the LBDS has a the responsibility of completing the protective task with a safe beam dump. As every operation always terminates with a beam dump, this system is the one for which safety has to be certified largely SIL3 or even better. The analysis of dependability of a simplified MPS will be given later in Chapter 8.

[8] More in general, the set of the interlocked systems depends on the operational phase of the LHC. The full set is necessary only for high energy and high intensity beam and, in the remaining period, a subset of the users can be masked.

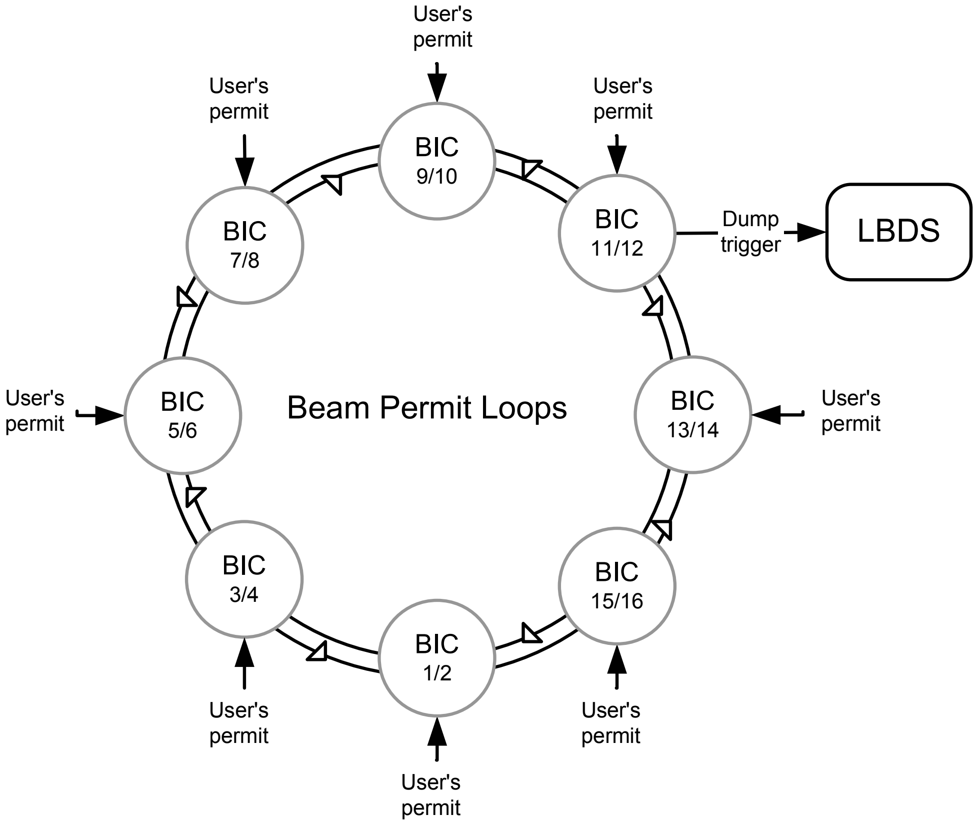


Figure 1.5: The Machine Protection System beam permit loop.

Chapter 2

The LHC Beam Dumping System

This chapter describe the functioning principles of the LHC Beam Dumping System and its components.

2.1 The System Inventory

The LHC Beam Dumping System (LBDS) has the role of extracting the beams on demand from the LHC rings and safely depositing them onto the absorbing block at the end of the dump channel. For each LHC beam the LBDS consists of the following components, see Figure 2.1:

- The MKD system is a series of 15 kicker magnet assemblies with their pulse generators that horizontally deflects the beam from the circulating orbit onto the extraction trajectory.

- The Q4 superconducting quadrupole is part of the optical elements of the circulating LHC beam but also enhances the horizontal deflection given by the MKD system to the extracted beam by more than 30%. It has an individual power converter.

- The MSD system is a series of 15 septum magnets for the vertical deflection. There are three types of MSD; the MSDA, MSDB and MSDC, which differ in septum thickness and their magnetic field. They are connected to a single power converter.

- The MKB system consists of a series of 10 kicker magnet assemblies and their pulse generators, arranged into 4 MKBH and 6 MKBV systems for the horizontal and vertical dilution of the beam respectively. Due to the dilution of the beam, the MKB magnets reduce the beam energy density when it arrives at the absorbing block.
- The TDE absorbing blocks are at the end of the beam dump lines. The extracted beam impacts onto the TDE graphite block stamping a characteristic 'e' twisted shape profile (15 × 25 cm), imposed by the MKB system, see Figure 2.2. This is important for avoiding any damage of the TDE due to overheating.
- The TCDS is a passive element that protects the MSD magnets from beam impact. The TCDQ and the TCS are passive elements that protect the Q4 and the downstream LHC from beam impact.

The system inventory also includes: FPGA based electronics to generate the synchronized triggering of the kicker magnets; PLC to guarantee the beam energy tracking of the power converters in the system, general status surveillance and diagnostics. The vacuum system of the beam pipes, the nitrogen over-pressure for the TDE and the different types of beam instrumentation are also part of the LBDS but are not shown in Figure 2.1. The main LBDS parameters are summarized in Table 2.1.

2.1.1 Operational Modes

Three different operational modes of the LBDS can be defined: the ready mode, the firing mode and the post-operational mode. The LBDS is required to be in the **ready mode** at any moment when there is beam in the LHC. The triggering system of the LBDS receives the revolution frequency from the RF

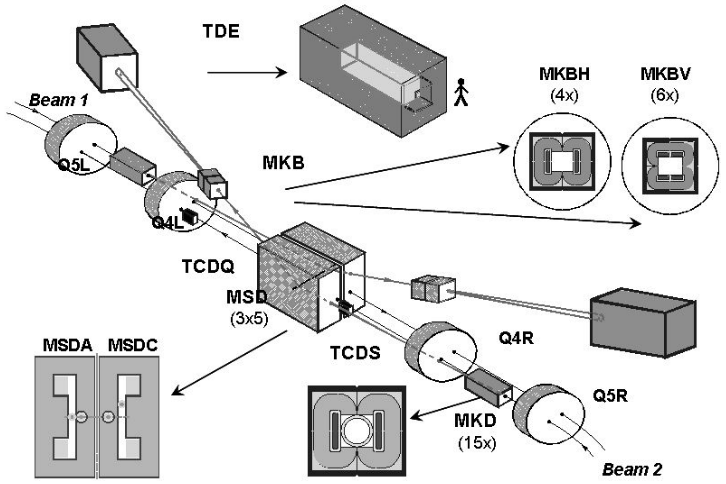


Figure 2.1: LBDS essential layout (courtesy of M. Gyr).

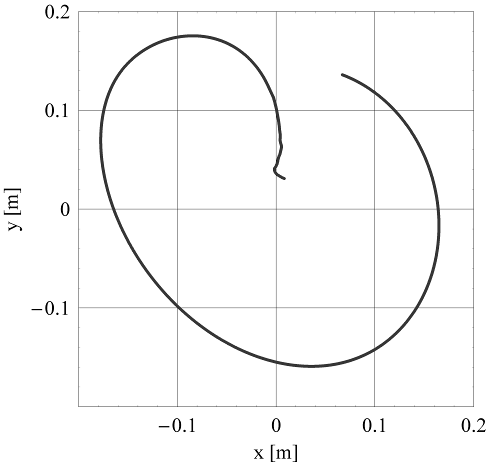


Figure 2.2: The twisted ‘e’ beam shape profile at the TDE target (courtesy of B. Goddard).

System	N. items	Length [m]	$\oint Bdl$ [Tm]	Deflection [mrad]
MKD	15	22.5	0.25	0.240
Q4	1	3	0.08	0.090
MSD	15	67.5	Type A: 0.8 Type B: 0.99 Type C: 1.17	2.400 (total)
MKBH	6	7.6	1.64	+/- 0.28
MKBV	4	7.2	1.077	+/- 0.28
TDE	1	8	-	-
Dump line	-	975	-	-

Table 2.1: Main LHC Beam Dumping System parameters.

system, which is phase-locked to the beam abort gap, a time interval of $3\mu s$ where the ring is deliberately left free of particles and corresponds to the rise time of the magnetic field of the extraction kicker magnets MKD. The beam energy is measured by the beam energy measurement system (BEMS) that translates the current measured at the power converters of the LHC main dipoles into a beam energy value, which is delivered to the MKD and MKB systems and applied as voltage settings of the local power converters. The voltage settings of the MSD and Q4 also track the beam energy, although they are generated in a different way, external to the LBDS.

As soon as a dump request is generated the system passes to the **firing mode**. The triggering system receives the dump request from the beam-interlocking controller BIC (see Chapter 8), which is converted into a trigger signal synchronously distributed to the MKD and MKB systems. As a result, the MKD kicker magnets will fire simultaneously and will all reach their nominal field in less than 3 μs. The nominal field is kept at least 90 μs, the time necessary for the removal of the entire beam from the ring. After the MKD the beam passes through the Q4 and the MSD magnets, reaching the MKB magnets for the dilution.

The system moves to the **post operational mode** after the beam dump. Currents and voltages from magnets generators, power switches, etc. are recorded and processed in post mortem diagnostics in order to check that everything has functioned as expected. In particular, they make it possible to discover faults that have either accumulated undetected during the operation, or occurred at the moment of the beam dump. If this check is passed successfully, the system is re-armed and the local beam permit signal is generated and sent to the Beam Interlocking System. In case of discovered anomalies during the post operational mode further investigation could be necessary, with a temporary stop of LHC operation.

2.2 The MKD System

The MKD system of the LBDS consists, per beam, of a series of 15 kicker magnet assemblies with their individual generator. The functional architecture of the MKD system is shown in Figure 2.3. Each **kicker magnet** is about 1.5 meter long and consists of a tape wound steel yoke with a one-turn winding at either side of the beam aperture, surrounded by a mechanical support frame. The beam aperture is delimited by a ceramic chamber that also acts as vacuum barrier with a vacuum of about 10^{-11} mbar. The magnet is connected to its generator by 8 parallel high voltage low inductance coaxial cables. Each **generator** consists of two identical redundant branches in parallel, charged to a voltage proportional to the beam energy signal that is received from the BEMS, see section 2.7. The high intensity current pulse of around 20 kA is the result of the discharge of a primary capacitor, charged up to 30 kV (depending on the beam energy), through a solid-state switch[1]. Two circuits compensate the overshoot, the first (OS1) charged at 350 V, and the second (OS2) beam energy tuned around 300 V. Both compensation circuits discharge through the same switch. All power switches receive the trigger command from the triggering system via two redundant power triggers. The capacitors are of the self-healing type that means that an internal short-circuits only leads to a small reduction of the total capacitance, which can be monitored. The magnet current pulse reaches a maximum of 18.5 kA for

[1] The solid state switch is more reliable than traditional gas switches [8].

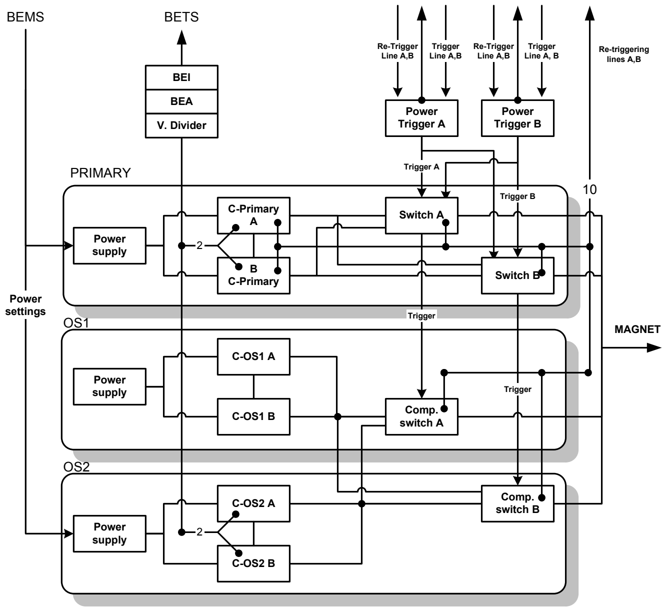


Figure 2.3: The functional architecture of one MKD generator assembly.

a 7 TeV beam, with less than 3 μs rise time followed by a period of at least 90 μs where the field varies by less than 7.5%. The expected beam deflection is shown in Figure 2.4.

Many failures of the MKD system are deemed catastrophic. The MKD system is designed fault tolerant and continuously surveyed in order to withstand these failures up to a certain limit or generating a failsafe operation abort once they are detected. These features are listed below:

- **Redundancy of the MKD systems**. 14 out of 15 systems are still able to perform a proper beam extraction[2].

- **Dual branch generator**. Each generator has two identical branches in parallel with an independent solid state switch in each branch. One branch may withstand the full current pulse in case of failure of the other.

[2] There exist few failure modes that overdo this redundancy. They are treated in section 6.3.

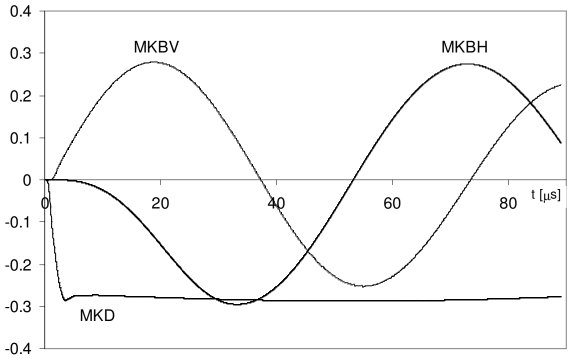


Figure 2.4: The beam deflection angles [mrad] for the MKD and the MKB(H/V) magnets (courtesy of J. Uythoven).

- **Redundant triggering**. Two independent power triggers drive the switches of both generator branches.

- **Surveillance of erratic triggers**. The re-triggering system monitors the erratic triggers in the MKD generators by 10 current pick-ups (placed at different points in the circuit) per system. In case of a detected erratic trigger, all MKD (and MKB) systems will be triggered asynchronously with the beam abort gap.

- **Surveillance of the beam energy tracking**. The primary and overshoot (OS2) capacitor voltage settings are monitored. The values are acquired by a local Beam Energy Acquisition (BEA) card and transmitted to the Beam Energy Interlocking (BEI) card where they are compared to the present beam energy. One BEA-BEI card per MKD magnet generator exists. A dump request is issued if the difference between the reference beam energy and the measured settings exceeds the 0.5%.

Many signals are recorded at the moment of the dump trigger like the currents in the power triggers, in the switches and at the magnets. Their analysis in post mortem diagnostics permits to discover failures that have accumulated silently in the system.

2.2.1 Power Triggers for the Kicker Magnets

The power trigger receives the trigger signal from the triggering and the re-triggering systems (sections 2.5 and 2.6) and re-transmits this signal, shaped and amplified, to the MKD or the MKB generator switches, see Figure 2.3. The system is housed in a dedicated VME crate, one per MKD system, and consists of two identical Power Trigger Modules (PTM) working in parallel, as shown in Figure 2.5. Each PTM consists of a primary driver that receives the signal from the triggering and re-triggering systems. The driver commands the power switch (three Insulated Gate Bipolar Transistors, IGBT, in series) for the discharge of a capacitor at a voltage that ranges between 800 V and 3000 V, as calculated by the power trigger controller (PTC) on the basis of the present beam energy. A mono-stable circuit generates the resulting output pulse. For reliability reasons, a second redundant path exists, which makes it possible that the trigger signal reaches the power switch even in case of failure of the primary driver[3]. A compensation driver commands another power switch (IGBT) for the discharge of an internal capacitor, which lengthens the pulse for the activation of the compensation circuit of the MKD generator. The compensation driver is designed to fire only if the primary has already fired, avoiding the dangerous scenario where only the compensation pulse is generated. Independent powering exists for the Power Trigger Controller (PTC-PS), the PTM (PTM-PS) and the primary capacitor of the trigger module (HV-PS). The resulting trigger pulse has the duration of 3 μs with 200 ns rise time and 400 A peak current.

The failure of the power trigger may lead to catastrophic consequences especially when it results in an erratic trigger for the MKD system. Fault tolerance and surveillance make it possible to withstands failures or issue an operation abort in the case that the failure is detected:

- **Dual branch trigger module**. The trigger signal is generated in two independent modules. One module is able to withstand the failure of

[3] In this case, the current flowing through the second path will damage the switch and the card, which will then need to be replaced.

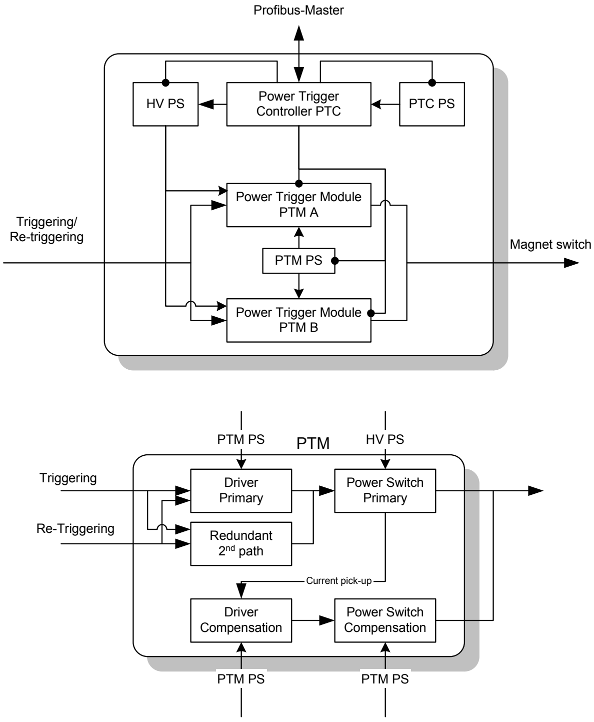


Figure 2.5: Functional architecture of the power trigger.

the other. As a resulting drawback the likelihood of erratic triggers is doubled.

- **Surveillance of power supplies failure**. The Power Trigger Controller (PTC) provides the continuous surveillance of the power supplies (+/-150V, 48V and 15V). The power supply of the primary capacitor within the Power Trigger Modules (PTM) is also monitored.

The input and output currents of the power triggers are monitored for post mortem diagnostics. They may reveal either a false contact or a missed trigger from an input line[4].

2.3 The MSD System

The MSD system consists of a series of 15 septum magnets comprising five MSDA, five MSDB and five MSDC. The functional architecture is shown in Figure 2.6. Each magnet is about 4.5 meter long and consists of a laminated iron-dominated frame, built using a welded construction of two half-cores, the coil and the septum. One chamber exists for the circulating beam and one for the extracted beam, kept at a vacuum of 10^{-11} mbar and 10^{-8} mbar respectively. One power converter supplies all magnets with a current that depends on the actual beam energy [9]. The nominal septum current for a 7 TeV beam is 880 A, resulting in an integrated magnetic field of 0.80 Tm in the MSDA, 1.0 Tm in the MSDB and 1.17 Tm in the MSDC magnet.

No failure of any magnet in the MSD can be tolerated and the consequences are severe. To reduce this risk, the system implements continuous surveillance:

- **Surveillance of the beam energy tracking**. The local BEI compares the measured current in the generator PC, which is identical to

[4] The perfect coverage of all failures is impossible. For instance, one major concern is the fault of the clamping diodes in the input section that leaves the module unprotected from over voltages with an increased risk of breakdowns.

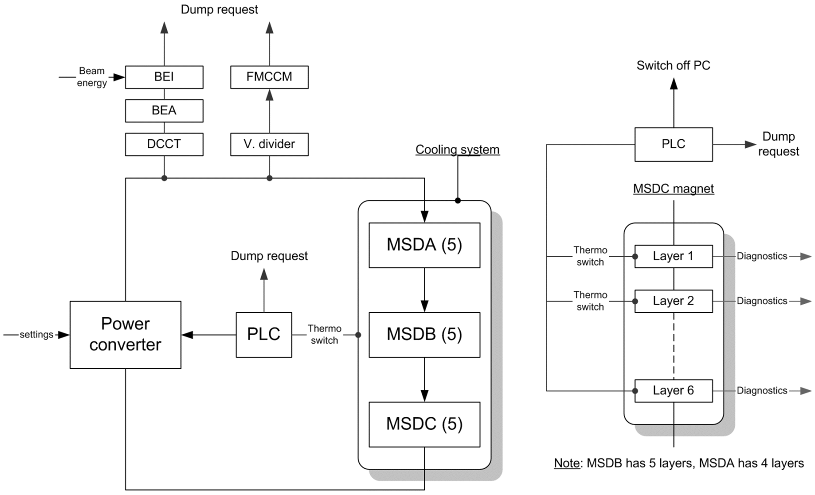


Figure 2.6: The functional architecture of the MSD system: the system (left) and one MSDC magnet (right).

the current in the magnets, to the theoretical value derived from the present beam energy. The dump request is generated and delivered to the BETS if a 0.5% error is detected (see section 2.8). This check is also effective for slow load changes that make the field change beyond the 0.5% tolerance[5].

- **Surveillance of fast magnet field decay**. A short in a magnet coil or in the PC is expected to provoke a fast field drop in less than 1 ms that is covered by the Fast Magnet Current Change Monitor (FMCCM) [96].

- **Surveillance of coil overheating**. The septa magnets are water-cooled. In case of a cooling system failure any overheating is detected by thermo-switches (one per coil layer, 75 in total for the MSD system) and an alarm will be sent to the local PLC that generates a dump

[5] As the septa are continuously powered the survey of their current is a better guarantee of their proper operation than the survey of the voltage of the pulsed magnets. Some of these failures are also caught by the PC internal surveillance, which is not included in the analysed system.

request and will subsequently switch off the PC.

Currents and voltages are collected for post mortem diagnostics in order to discover failures that have accumulated silently, as well as slow drifts in the electrical parameters that lead to a degraded magnet field.

2.4 The MKB System

The MKB system of the LBDS consists of 10 kicker magnets arranged into 4 MKBH, for the deflection in the horizontal plane, and 6 MKBV, for the deflection in the vertical plane. The magnet length is 1.9 m for the MKBH and 1.2 m for the MKBV. The functional architecture is shown in Figure 2.7. The system is mounted in a vacuum tank with a pressure of 10^{-6} mbar and no vacuum chamber for the beam is required. Each magnet is connected to its generator by 10 parallel low inductance high voltage coaxial cables. Each generator is powered to a voltage proportional to the calculated beam energy and triggered by the triggering system. More in detail, the generator for the MKBH consists of an oscillating capacitor circuit CH (pre-charged up to 16.4 kV) that discharges through a solid-state switch[6] (switch-H). The resulting current pulse is an attenuated sinusoid, 25 kA of amplitude and 70 μs of period, which at its maximum causes a beam deflection of 0.278 mrad in the horizontal plane for the total of four MKBH magnets, see Figure 2.4. Similarly, the MKBV consists of one resonant capacitor circuit CV (pre-charged up to 22.3 kV) that discharges through a solid-state switch (switch-V). The resulting current pulse is an attenuated sinusoid, phase shifted by 90 degrees with respect to the MKBH pulse, which produces a beam deflection of 0.277 mrad in the vertical plane for the six magnets, see Figure 2.4.

The failure of the MKB system is severe only if all magnets are lost for the vertical or the horizontal dilution, which is a conservative definition. In all other cases, dilution of the beam energy is reduced but still acceptable,

[6] The solid-state switches are identical to those used in the MKD generators and receive their trigger signal from one power trigger per magnet.

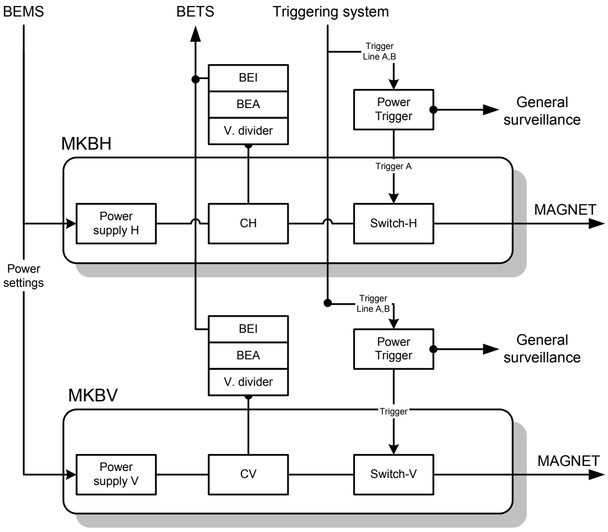


Figure 2.7: The functional architecture of one MKB generator.

resulting in possible overheating of the TDE block. Differently from the MKD system, the architecture of the generators is not redundant. There is only one power trigger instead of two and one branch instead of two per generator. For this reason, the system has continuous surveillance for the energy tracking failures with one BEI card per magnet generator, connected to the BETS. During operation, the currents at the local power triggers, the power switch and at the magnet are collected for post mortem diagnostics.

2.5 The Triggering System

The triggering system delivers the synchronized trigger signal to the power triggers of the kicker magnet generators [4, 16]. It consists of two independent VME (LinxOS) boards housed in one VME (64X) crate. Each board is connected to the local Beam Interlock Controller (BIC point 6, see Chapter 8) via an optical interface. Dump requests are received and stored into a buffer. The signal that drives the buffer is generated by an oscillator implemented in an FPGA, which keeps it tuned to the beam revolution frequency and locks the phase to the beam free gap. Once the dump request

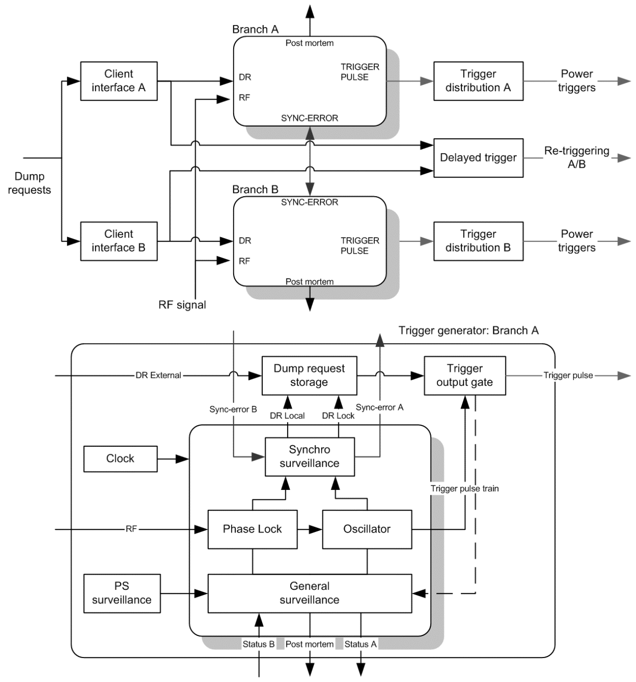


Figure 2.8: The functional architecture of the triggering system.

has been received, this is transmitted (synchronized) to an output trigger gate where it is shaped and amplified. A fan-out current transformer distributes the output pulse to the power triggers of the MKD and MKB kicker magnets. For reliability reasons, a delayed ($> 90\mu$s) asynchronous trigger is also sent to the re-triggering system [17, 77]. The functional architecture of the triggering system is shown in Figure 2.8.

The failure of the triggering system is catastrophic only in case the trigger is not transmitted to all magnets (i.e. generation part failure) or it is transmitted to less than 14 MKD (i.e. fan-out distribution failure). The synchronization error is not critical due to the passive protection elements TCDQ, TCDS and TCS that minimize the consequences of a sweep of the beam over the Q4 and MSD magnets. However it is unwanted because of the likelihood of generating quenches of the superconducting magnets, and the increased risk of damage if, for example, the TCDQ is not in its correct position. To reduce the likelihood of these failures, the system is fault tolerant and continuously surveyed.

- **Redundancy in the trigger generation part**. The trigger generation is made of two redundant and identical modules so that one trigger generation failure can be tolerated.

- **Redundancy in the trigger distribution**. The trigger distribution consists of two independent paths for the synchronized trigger so that one distribution failure can be tolerated. The re-triggering lines still carry a delayed asynchronous trigger in case of the complete failure of the trigger distribution.

- **Surveillance of synchronization failures**. A crosschecking mechanism detects local and external synchronization errors with the beam free gap. Two internal synch-error bits are generated within the triggering system FPGA and shared between the trigger generators A and B. When a failure is detected, the status of one of these bits changes. The failed trigger generator must inhibit its dump request buffer and communicates its failure to the other trigger generator that issues a local dump trigger [77].

The output currents of the trigger generators and the trigger distribution lines, as well as the FPGA status, are recoded for post mortem diagnostics.

2.5.1 The VME Crate

The VME crates house the electronics boards of the triggering system and other LBDS electronics. The crate consists of one power supply module, one fan-tray module and the VME backplane designed to accept user boards that comply with the VME bus standard (6U × 160 mm). The power supply module receives 230 V AC main power from the UPS unit, which is internally converted into 15 V, 5 V and 3.3 V (DC). The system is cooled by a fan-tray consisting of three fans plus the fan control unit, which governs the rotation speed of each fan, depending on the air temperature. Internal surveillance gives information on the status of the three converters and a dump request is generated in case a powering failure is detected. A beam dump request can also be programmed in case of failures of a fan. Customizing the VME crate could reduce the contribution to the system unavailability. For instance, one fan failure might be tolerated without generating a dump request and a second power supply module could be added in parallel to reduce the VME failure rate.

2.6 The Re-triggering System

The re-triggering system is designed to catch erratic triggers in the MKD system and re-distribute them to all 15 MKD generators. If one MKD kicker fires, the other kickers will be triggered with a maximum delay of 700 ns for beam energies above 3 GeV/c. This action is not synchronized with the beam abort gap and produces beam losses to the septa and the arc aperture. Again, these losses are intercepted by the TCDS and TCDQ/TCS. The functional architecture is shown in Figure 2.9. The system consists of two independent re-triggering distribution lines A and B connected to the MKD power triggers via two Re-Triggering Boxes (RTB) [16, 17]. The re-triggering line also receives a trigger signal from the triggering system. Each re-triggering box picks up the current at different points in the primary and secondary circuit of the MKD pulse generators. It is connected to the power triggers A and B of the local MKD generators and to the re-triggering line. Once the erratic trigger reaches the RTB, it is distributed to the other RTB by a domino effect, using the energy stored at each stage.

The failure of the re-triggering system leaves the trigger event uncovered in all or part of the MKD systems. The re-triggering system is not surveyed[7] but it implements redundancy in order to withstand failures at a reasonable extent.

[7] The re-triggering system is realized in passive components that make it unable to generate spurious triggers.

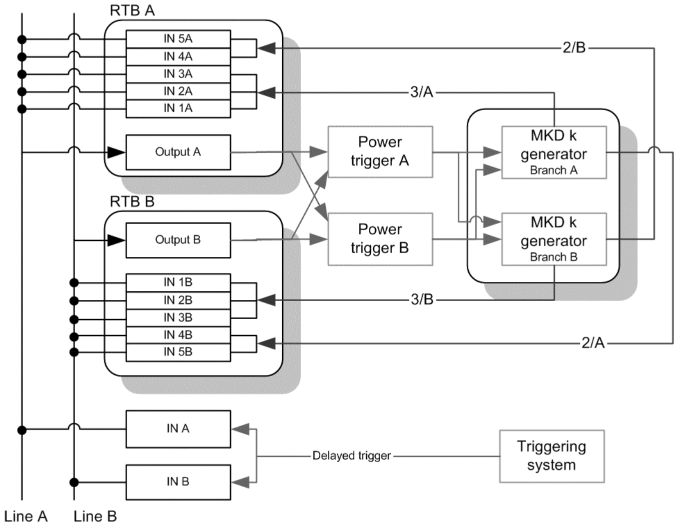


Figure 2.9: The functional architecture of the re-triggering system.

- **Redundancy in the current measurements**. Every current in the circuits is measured twice, and go to different RTBs, which guarantees a higher fault tolerance (see Table 2.2). For example, the primary capacitor current is picked up by one channel to the Re-Trigger Box A and one to the Re-Trigger Box B. This also assures the coverage of the erratic trigger at the power trigger, upstream the MKD generator branches.

- **Double triggering lines**. The re-triggering lines are doubled in order to withstand the failure of one of them.

At every beam dump the re-trigger signals (currents) are distributed to the kickers, which are analysed by post mortem.

2.7 The Beam Energy Measurement System

The Beam Energy Measurement System (BEMS) calculates the beam energy that is distributed to the MKD and MKB systems in order to derive their settings [16]. The system receives the measured current of the LHC

Input channel	**Source of erratic**	**Power trigger**	**MKD primary**	**MKD OS1/2**
Re-triggering line A				
IN_{A1}	Primary capacitor A	A, B	A, B	
IN_{A2}	Primary switch A	A, B		
IN_{A3}	Comp. switch A	A, B		A
IN_{A4}	Primary switch B	A, B		
IN_{A5}	Comp. switch B	A, B		B
Re-triggering line B				
IN_{B1}	Primary capacitor B	A, B	A, B	
IN_{B2}	Primary switch B	A, B		
IN_{B3}	Comp. switch B	A, B		B
IN_{A4}	Primary switch A	A, B		
IN_{A5}	Comp. switch A	A, B		A

Table 2.2: Coverage of the erratic trigger events.

main dipole power converters at the points 4/5 and 7/8. To improve the reliability and for internal data validation, each current is measured twice by two Direct Current-Current Transformers (DCCT), connected to one Beam Energy Acquisition cards (BEA), as shown in Figure 2.10. The BEA acts as an Analogue to Digital Converter (ADC) between the power converter of the dipole magnets and the BEMS. A multiplexer alternates the input to the ADC module between four analogue values: two reference voltages and the two measurements from the DCCT of the same dipole magnet. The four values are converted in the ADC into 16-bit digital values. A noise filtering is applied by calculating the average of 16 samples for each of the 4 inputs. Results are encoded and transmitted through a serial optical link[8] to the BEM card where they are decoded and treated separately. In this phase possible transmission, reception and timing errors will be detected. If no errors are detected, the four values are sent to two voters for the internal crosscheck: the input 1 is compared to the input 3 and the input 2 is compared to the input 4. The four values are then averaged and converted into a beam energy value, using a pre-loaded look up table saved into a flash-ROM and moved to two internal ROMs of the FPGA. The BEM sends the calculated beam energy to the the MKD and MKB magnets via an 8 bits bus where it is

[8] A Cyclic Redundancy Check (CRC) and Manchester encoding are used.

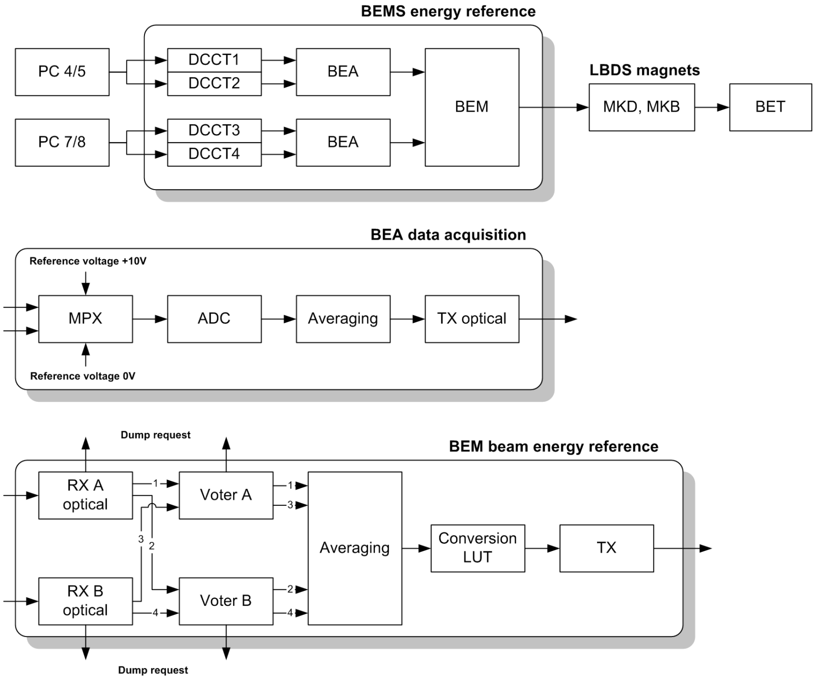


Figure 2.10: The functional architecture of the BEMS.

finally translated into voltage settings.

The failure of the BEMS can be catastrophic due to the fact that it may deliver a wrong beam energy reference to the magnet power converters. In order to reduce this risk, the BEMS implements on-line surveillance over the the data processing for many input and output quantities. Part of the failures generated in the system are also detectable by the BETS.

- **Redundancy in the data acquisition**. The LHC dipole magnet currents used to determine the beam energy are measured at two different sources at point 4/5 and 7/8. At each point the measurements are taken by two DCCTs.

- **Surveillance of Transmission/Reception errors at the BEA-BEM interface**. A dump request is issued in case the BEM has received a corrupt packet.

- **Surveillance of timing errors in the BEM**. A watchdog timer surveys the normal flux of operations. Timing errors with a resolution of 1 ms are caught and a dump request is generated.

- **Surveillance of wrong values before conversion**. A voting mechanism checks that the values received from the two BEA cards are in agreement. A dump request is issued in case of any inconsistency.

During operation, various signals are collected for post mortem diagnostics in order to discover failures that have accumulated silent in the system, FPGA included. There still remains subtle failures that are not detectable: if the current reading of all dipole PCs is erroneous and if the look-up tables have identical faults for the two BEMS. Those kinds of failures are called systematic or common mode type. Nevertheless, as they can only happen at the system start up, they are discovered during the dumping of the low intensity pilot beam in the operational mode, see section 2.1.1, with no resulting damage to the machine.

2.8 The Beam Energy Tracking System

The Beam Energy Tracking System (BETS) continuously checks that the settings in the LBDS power converters agree with the actual beam energy within a ±0.5% error tolerance [16]. The architecture is shown in Figure 2.11. The voltages of the MKD, MKB generators and the currents of the MSD and Q4 generators are acquired via BEA-BEI cards. In total, there are 27 BEA-BEI per beam, 15 for the MKD, 10 for the MKB, 1 for the MSD and 1 for the Q4. The BEI interfaces to the BEA with an optical receiver. The received values are first decoded then averaged and converted into beam energy values according to an internal look up table. As a final step, a voter compares the calculated energy value to the value received on the VME bus from another BEMS, which takes the measurements at the dipoles 5/6 and 6/7. A dump request is issued if the difference between any pair exceeds the preset ±0.5% threshold. The dump request mechanism is implemented in

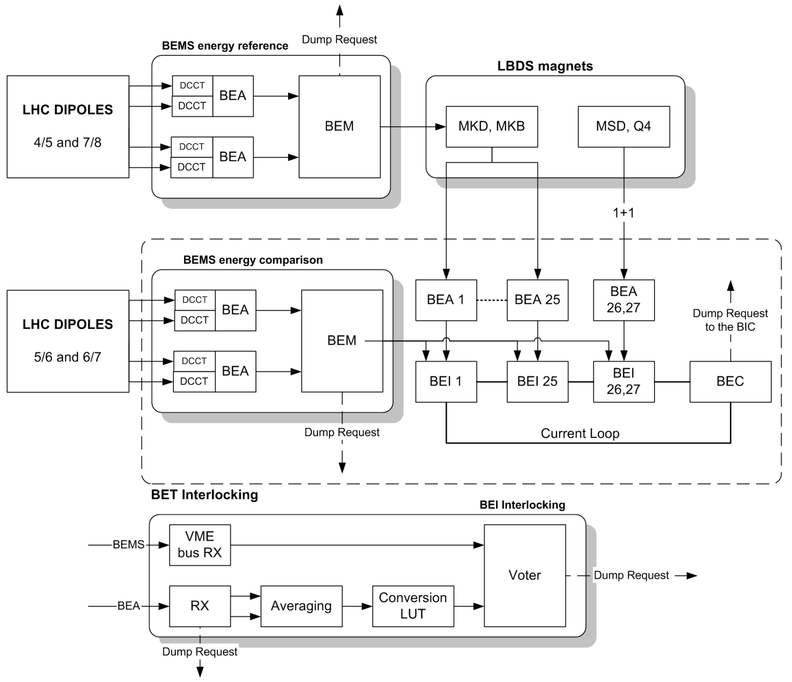


Figure 2.11: The functional architecture of the BETS.

Data acquisition channels	**BEMS**	**MKD**	**MKB**	**MSD**	**Q4**
BEA-BEI 1 . . . 15	All set	1 per magnet			
BEA-BEI 16 . . . 25	All set		1 per magnet		
BEA-BEI 26	Yes			Yes	
BEA-BEI 27	Yes				Yes

Table 2.3: BETS coverage matrix

the BETS with a current loop. Each BEI is able to cut the loop with the generation of the local dump request. This event is detected by the Beam Energy Controller (BEC) and transmitted to the triggering interface and the Beam Interlock Controller of the machine protection system. Like the RTS, the BETS is able to cover the energy tracking failures distributed in the LBDS, as shown in the coverage matrix of Table 2.3. The systems runs continuously self-surveillance over the whole data-processing, in the BEMS, in the BEI-BEA cards, and in the BEC.

A failure of the BETS leaves the LBDS uncovered with respect to all powering failures or just a part of them, as it is shown in the coverage matrix of

Table 2.3. One BEI card per MKD and MKB generator assures the coverage of the energy tracking failures, one is for the MSD power converters while the failure in the BEMS is covered by the entire set of BEI cards. The BEC is responsible for the management of the dump requests from all BEI cards and represents a single point of failure for the system. In order to reduce the likelihood of these failures, the system verifies continuously the correct functioning of its parts. After a dump request, the BET system is checked with post mortem diagnostics and potential hidden failures are discovered.

Chapter 3

Introduction to Dependability

This chapter introduces to dependability engineering with a special attention to the architectures used for safety critical applications.

3.1 The System Dependability Attributes

Dependability is the measure for the quality of service in time given by the system. It encompasses the notions of availability, reliability, safety, maintainability and other more specialized attributes. In [53] dependability is defined 'the ability of the system to deliver a service that can be justifiably trusted' but other definitions are given by international standards authorities like ISO. Definitions of some dependability attributes are:

- **Availability**: the readiness for correct service.
- **Reliability**: the continuity of correct service.
- **Safety**: absence of catastrophic consequences in case of failure.
- **Maintainability**: the ability to undergo modifications and repairs.

Availability distinguishes from reliability for the possibility of withstanding more service outages during the system lifetime, just one outage being unacceptable for a reliable system. Safety distinguishes from availability and reliability for the consequence of the service outage, which is ranked according to a severity level. Maintainability is the measure of the repair process including fault diagnosis, localization and isolation plus repair or replacement [2]. The dependability attributes can be also combined with other quantities (e.g. costs, quality of service, etc.) into indexes of performance, classified in literature under the term performability [88].

3.2 The Failure Process

The system failure is the final pathology of the failures of its components and their propagation. Its description is based on the fault-error-failure model, also called the 'chain of threats' [3]. Faults occurring at physical level are activated by patterns that can be reproducible (i.e. hard faults) or not (i.e. soft faults). The reproducible faults are also called permanent faults and move the component into a persistent faulty state. Faults are called transient if they happen under certain conditions that are difficult to reproduce and predict. Failures are called random if they occur due to progressive degradation (hardware), or systematic (hardware and software) if they are introduced during the system life cycle [46]. Faults can develop independently but there may exist causes that provoke them simultaneously, generating a common mode failure, which is the most undesired event for the system. A detailed classification of faults and failures can be found in [53] and [73].

The dynamics of failures are intrinsically complex to analyse because of their random nature and the dimension of the failure space, often larger than the space of admissible states. A qualitative analysis, consisting of the classification the system failure modes, is necessary before looking at their likelihood through a quantitative analysis.

A **qualitative analysis** enumerates all failure mechanisms in the system and their consequences. Failure Modes and Effects Analysis (FMEA) or Failure Modes Effects and Criticalities Analysis (FMECA), if criticality (C) is also a concern, are systematic techniques for qualitative failure analysis [39, 65]. The FME(C)A studies all fault-error-failure chains in the system, which are classified by name, the way they occur and propagate, the possibility of being compensated and/or detected, the consequence and their criticality.

A **quantitative analysis** provides a statistics of the Time To Failure (TTF) of a component. This is the hazard function, which is the probability that the component will fail at time t given it has not failed before [39]:

$$h(t) = \frac{f(t)}{R(t)} = -\frac{dR(t)/dt}{R(t)} \tag{3.1}$$

where $R(t) = 1 - \int_0^t f(\tau)d\tau = exp[-\int_0^t \lambda(\tau)d\tau]$ is the reliability, $f(t)$ is the probability density function of TTF. For hardware components, the hazard function experiences (in time) the characteristic bathtub curve with three distinct periods: the burn-in, the useful and the wear-out period [2], see Figure 3.1. The burn-in period is characterized by 'infant mortality' mainly due to production or manufacturing errors. As soon as these are discovered and eliminated, the hazard function decreases and a Decreasing Failure Rate (DFR) is expected. A Constant Failure Rate (CFR) characterizes the useful period. During the wear-out period aging and wearing processes dominate the system, which suffers from an Increasing Failure Rate (IFR)[1].

The probability distribution of the TTF must account for the failure during the full lifetime of the component. The Weibull distribution is one of these. It is characterized by two parameters[2], the shape factor α and the scale factor λ [39]:

$$F(t) = P(T \leq t) = 1 - \exp[-(\lambda t)^{\alpha}] \qquad t \geq 0 \tag{3.2}$$

The average or Mean TTF (MTTF) is $E[T] = \Gamma(\frac{1}{\alpha} + 1)$ and the variance is $Var[T] = \frac{1}{\lambda^2}[\Gamma(\frac{2}{\alpha}+1) - \Gamma^2(\frac{1}{\alpha}+1)]$, where the function Γ is the Gamma function[3]. The hazard function $h(t)$ is $\alpha\lambda(\lambda t)^{\alpha-1}$. By choosing the parameters α

[1] The hazard function for software does not have the wear out period and it is characterized by a long burn-in period due to the removal of systematic failures.

[2] The Weibull distribution is usually referred in literature as Weibull[α,1/λ].

[3] The gamma function is $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx$, $\alpha > 0$. The gamma density function is $f(t) = \lambda^\alpha t^{\alpha-1}e^{-\lambda t}\Gamma(\alpha)^{-1}$.

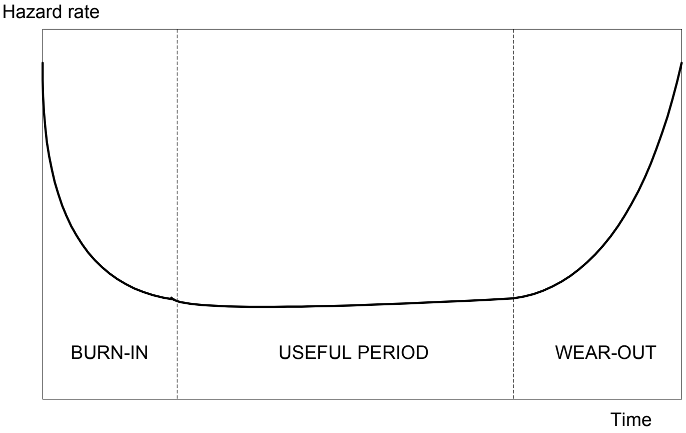


Figure 3.1: The bathtub curve of the hazard rate function.

is possible to obtain DFR ($\alpha < 1$) or IFR ($\alpha > 1$). For $\alpha = 1$, the Weibull has constant failure rate λ and corresponds to the exponential distribution, which is the most used distribution for the description of random failure processes. The MTTF of the exponential distribution is $1/\lambda$ and is defined by the component surviving at t = MTTF with a probability of 0.37. The standard deviation of the exponential distribution is $1/\lambda$. An introduction to the families of TTF distributions can be found in [83] and [39].

The TTF distribution of a component is estimated from failure reporting, during the system use, or from reliability runs and accelerated life testing [2, 83]. As an alternative, the TTF statistics are calculated by reliability prediction using existing literature on the components failure rates. The reliability prediction is a mathematical tool and it is cheaper (in time and money) with respect to testing, though less accurate. The most popular reliability prediction tool is the Military Handbook 217 (MIL-HDBK 217), issued in many versions dating from the early sixties by the Department of Defense of the USA [66] and recently by the Reliability Analysis Center (RAC) [31].

The MIL-HDBK 217 provides two different methods for the calculation of failure rates: the stress analysis and the parts count analysis [66]. The **stress analysis** is the most complete one. The failure rate is calculated as function of various parameters: environmental and operational temperatures, humidity, electrical fields, vibrations, radiations, voltage and current ratings, power and quality [71]. The failure mechanism is activated according to the Arrenius's equation $TTF = Ce^{E_A kT}$, where E_A is the activation energy of the component failure, k is the Boltzmann constant and T is the temperature, The Arrenius's equation expresses the development of a failure by analogy with a chemical reaction [83]. The **part count analysis** is a simplified version of the stress analysis. A base failure a rate λ_b is calculated with the stress analysis for standard operating conditions, then it is adjusted with respect to the quality factor and the environment, that is $\lambda = \pi_Q \lambda_b (\pi_E)$, some 14 different environments existing from ground benign to airborne critical [12]. The parts count analysis is useful in the early design phase when the necessary information is either lacking or unreliable. Whatever the used method, the calculated failure rates are defined as the 90% confidence interval estimate.

Reliability prediction does not apportion the failure rates into failure modes. This becomes very important when one is concerned with the consequences of a failure. A method for the apportionment of failure modes is provided by the RAC FMD-97 [34] and in part also by the military standard MIL-HDBK 338B [67] for a large set of components.

Other methods for reliability prediction exist. Some of these are specialized to the telecommunications like the Telcordia of Bell and AT&T industries, the Siemens and CNET of France Telecom. The IEC 62380 [43] of the International Electro-technical Commission (IEC) is derived from the MIL-HDBK 217. Benefits and drawbacks of the different methods compensate each other, depending on the field of application and the aim of the analysis, so that it is difficult to say a-priori what is the best [48]. The true alternative to the reliability prediction is the physics to failure approach [91]. The mechanisms of failure are more accurately defined but the philosophy that is behind is totally different. No failure rates are provided but guidelines that help to prevent flaws in the component production, those potentially impair-ing the TTF, i.e. more quality assurance in the production stage rather than reliability prediction.

Many criticisms exist about the reliability prediction using the MIL-HDBK 217: the assumption of constant failure rate bounds the applicability of the tool to the useful period, which for many components ranges between 2 and 5 years only [30]; the failure rate is not apportioned in failure modes [12]; the failure rates database quickly becomes obsolete, especially for electronic technology that changes in few years; the stress analysis based on the Arrenius's equation is someway arbitrary [91].

Despite these criticisms, the reliability prediction is still the most used tool to assess the component reliability. Many complex systems have been certified using the MIL-HDBK 217 like the international space station ISS, civil and military airplanes, avionics systems, nuclear and chemical plants and so on. As opportunely remarked by the RAC president J. Fedduccia in [70]; the reliability prediction MIL-HDBK 217 is mostly a tool for comparing design alternatives, discovering weak points and bottlenecks rather than for obtaining precise reliability estimates, for which its results are inaccurate and possibly conservative. The tool is also not recommended as guideline for reliability improvement. For example, designing a costly cooling system for lowering the operating temperature and therefore the failure rate, or using ceramic packaging, more reliable but even heavier than the plastic ones, would lead to benefits that are doubtable and likely disproportionate to the additional costs. For analogous reasons, the simpler parts count method is usually preferred to the stress analysis method, which can be used to refine the analysis of critical components operating in extreme conditions. But, if this is the case, only testing will give the final response.

3.3 Dependability Modeling Techniques

Dependability modeling techniques aim at building a mathematical model for the description of the dependability attributes of the system as a function of the system architecture and the failure modes with their statistics, as derived by FME(C)A and reliability prediction. The analysis is performed for an assumed mission profile (i.e. the operational scenario), either analytically or by simulation, and returns a probability for the dependability attribute or a derived statistic[4]. The dependability modeling techniques split into two main categories: combinatorial and state based.

3.3.1 Combinatorial-based Techniques

Combinatorial techniques describe the system failure as the logic combination of failures occurring in its components. At the basis there is the definition of structure function of the system [39, 80, 83]:

Definition 3.1. The **structure function** $\Phi_A : X \rightarrow [0, 1]$ is a function of the state of the components X of the system, arranged with respect to the system architecture A, which returns the values 0 if failed or 1 if functioning.

The definition of the structure function may be applied at a lower level provided that the each component is given a binary variable x: 1 if functioning and 0 if failed. As an example, a non-fault tolerant architecture will be sensitive to the failure of every component, which is the case of the series architecture. On the contrary, a redundant architecture will be sensitive to the accumulation of failures, which is the case of the parallel architecture.

For any system with the architecture A, the structure function will always be between the series (S) and the parallel (P) architectures so that:

$$\Phi_S(X) \leq \Phi_A(X) \leq \Phi_P(X). \tag{3.3}$$

The structure function can be expanded in minimal terms: the minimal path sets ρ and the minimal cut sets χ [39].

Definition 3.2. A **minimal path set** is the set of components that are all necessary for the system to function.

[4] Statistics are the MTTF, the MTBF (Mean Time Between Failures), the MTTD (Mean Time To Detection) and the MTTR (Mean Time To Repair).

For a series system there is only one minimal path set ρ, which is the entire X, while in a parallel system $|X|$ path sets exist[5]. The structure function is:

$$\Phi_A(X) = 1 - \prod_k (1 - \rho_k) \tag{3.4}$$

and at least one path must be satisfied in order that $\Phi_A(X) = 1$.

Definition 3.3. A **minimal cut set** is a set of components, whose failure leads to the failure of the system.

The definition of cut set is complementary to that one of path set. In a parallel structure there is only one cut-set χ, which is X, while $|X|$ cut-sets exist for a series system. The structure function is:

$$\Phi_A(X) = \prod_k \chi_k \tag{3.5}$$

and all the cut-sets must be satisfied in order that $\Phi_A(X) = 1$.

The structure function is limited in its applications. It relies on the assumption that each component has only one failure mode, which is suitable for reliability calculations but not for safety. In addition, the translation of a Boolean expression into a probability is a delicate passage that is simplified under the assumption of statistic independence of failures.

Fault-trees and the reliability blocks are examples of combinatorial techniques [39]. **Fault-trees** describe the system failure arranged into a tree-like structure, where the root represents the system that has failed and the leaves account for the failures at the components level. A fault-tree calculates Prob$\{1-\Phi_A(X)\}$ in minimal cut sets. The **reliability blocks** describe the condition that the system is functioning as any fault-free path existing from two points, the source and the sink, in between of which there are the blocks, i.e. the system modules, arranged likewise the functional architecture. A reliability block diagram calculates Prob$\{\Phi_A(X)\}$ in minimal path sets. Examples of fault tree and reliability block diagram are given in Figure

[5] $|X|$ is the cardinality of the set X.

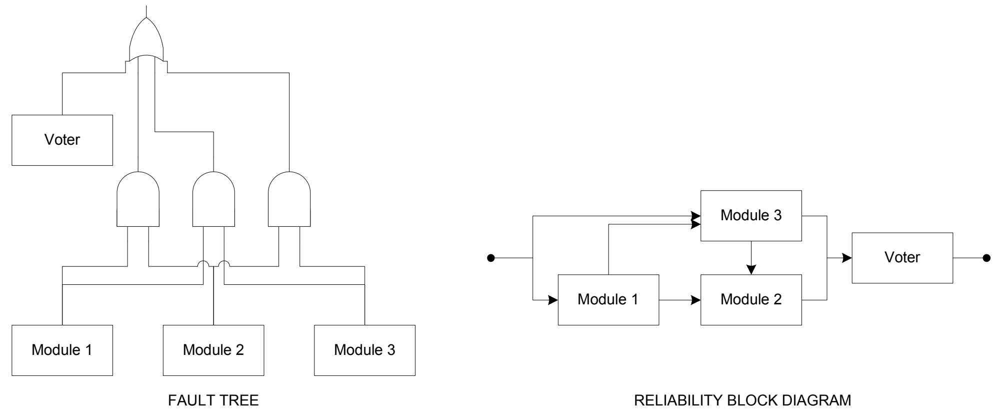


Figure 3.2: Fault tree (left) and reliability block diagram (right) of a Triple Modular Redundancy.

3.2 for the Triple Modular Redundancy architecture with voter presented in 3.4.1. The voter is assumed to be single point of failure[6]. The fault tree in Figure 3.2 (left) describes the logic condition of failure for the system, which is $\Phi_{TMR} = x_V(x_1x_2 + x_1x_3 + x_2x_3)$, where x = 1 if the component is functioning, 0 if failed. The reliability block describes the condition for the success of the system, which is $\Phi_{TMR} = 1-(1-x_V)(1-x_1)(1-x_2)(1-x_3)$.

The advantage of the combinatorial approach, and of fault trees in particular, is that they arrange the failure modes into a logical hierarchy, which is also useful as a database for diagnostics and failure reporting. On the contrary, the main limit of combinatorial techniques is the difficulty of modeling the cause-effect dependence between failures, because this would break through the assumption of statistical independence. The combinatorial techniques are ranked at the lowest level of the dependability modeling techniques [62]. Recent studies have tried to extend the modeling capability and the dynamic fault trees are an example [29]. Vast literature exists on this subject [39, 68, 84, 83].

[6] This is the case when the voter produces the average from the output three modules.

3.3.2 State-based Modeling Techniques

The paradigm of the Discrete Events Systems (**DES**) is suited to describe failure processes and dependability problems in general, assuming that they are governed by discrete events throughout a finite set of states.

Definition 3.4. A DES model is a stochastic timed automata, which is formally defined with a 5-tuple $\{X,\ E,\ \delta(x,e),\ x_0,\ \Psi\}$, where X is the space of states, E is the space of events, $\delta(x,e)$ is the state function that calculates the state transition given the event e has occurred, x_0 is the initial state and Ψ describes the distribution functions for the events in E [18]. The probability distribution of X at time t describes the evolution of the stochastic process.

The mathematics underlying the state-based techniques is more complex than the mathematics of the combinatorial techniques. Some assumptions are necessary in order to make the problem analytically treatable like for the Markov processes which will be introduced in section 4.3. Even when the solution is mathematically feasible, some millions states that can be generated for realistic case studies. In literature, this is called the states explosion problem. There are modeling languages that automatically generate the state transition diagram from a high-level description of the system behavior, as it is the case of the stochastic Petri nets. The model generated by the Petri nets is a DES that under certain conditions underlies a Markov chain [25]. Recent studies exist for deriving a Petri net from the UML (Unified Modeling Language) representation of the system [50].

The state-based approach is the most powerful in the modeling hierarchy [62]. It is suited to model a larger set of features that characterize a failure process, like cause-effect mechanisms, events concurrency, fault detection mechanisms, periodical inspections, etc. As rule of thumb, the state based approach must be preferred for failure processes that are observable and controllable in their dynamic. In absence of this, a combinatorial approach is usually preferred. Hybrid approaches exist in literature, which combine the advantages of state based and combinatorial techniques. Logic expressions (e.g. like in a fault tree) drive transitions between set of states [10, 11], with an expected reduction for the size of the problem.

3.4 Design for Dependability

Design for dependability can be considered the combination of various activities involving the different phases of the system lifetime[7] from specification and design passing through the prototyping, the manufacturing and the production, up to the installation and the operational phase [2, 49, 84]. These activities are classified in fault prevention, fault tolerance, fault removal and fault forecasting[8] [53]. Fault prevention avoids failure in the design phase by the use of quality assurance and system engineering procedures. Fault removal consists of checking the system functionality by formal methods (exhaustive in the exploration of the state space of the system) or testing via injection of fault patterns. These two activities are not treated here. Fault forecasting estimates the dependability attributes on a mathematical model of the system and is the subject of a next chapter. This section aims at introducing to fault tolerance and its general features.

Fault tolerance increases dependability by designing a system that is able to deliver the service, or some degraded version of it, even in the case of some specific classes of faults [46]. This is a collection of techniques that aims at reducing the likelihood of failure and preserving as long as possible the system functionality. The core of fault tolerance is redundancy by fault masking. The system has the possibility to withstand one or more faults without interrupting the service. In addition to this, fault tolerance includes

[7] Several standards have tried to incorporate the dependability design along the system lifetime, like for example the BSI standard 5760 [39].

[8] These four activities globally aims at making the system resilient to failures, which does not mean that a system, which is certified dependable, is complying with the functional specification. These two points of view, design for the function and design for the dependability of the function, have to be kept separated and sometimes they are a source of misunderstanding.

fault containment, which is the ability of confining the fault and to avoid the propagation, and fault recovery, which is the ability of treating a fault by detection, insulation and repair, before this might affect the system's functionality. The harmonization of these techniques, used together, is often managed by complex systems called fault management systems.

The type of fault tolerance to implement in the design of the system is also based on the results obtained by the FME(C)A, reliability prediction and fault forecasting in general [73]. Moreover, it depends on the dependability attribute. For example, design diversity[9] is introduced as a fault tolerant solution for common mode failures [1]. Redundancy and fault masking apply to reliability applications, fault recovery techniques is for availability applications and real-time fault management is for safety critical applications [85].

The elimination of failure in a fault tolerant system is feasible up to a certain extent only, because the complete elimination is either impossible or costly. The above considerations do not take into account the cost of the design. To include this would mean to make the fault tolerant design a problem of optimal apportionment of resources [82].

3.4.1 Static Fault Tolerance

Static (passive) fault tolerance implements fault-masking strategies [74]. The simplest scheme is the parallel redundancy where failures are tolerated up to a certain extent. Every time the system must deliver a unique value from the parallel computation, a voter of the outputs of the redundant components is needed. An example is the Triple Modular Redundancy (**TMR**), see Figure 3.3. In a TMR the voting mechanism is majority based (e.g., 2 out of 3), which permits the system to withstand one failure, but more sophisticated voting techniques exist depending on the delivered output. For example, average or median voters are used if the output is a signal to a controlled process (e.g., the command to the actuators). In this case the voter is

[9] Design diversity is redundancy for systems that provide the same function but implement different technologies, which reduces the possibility of sharing failures modes.

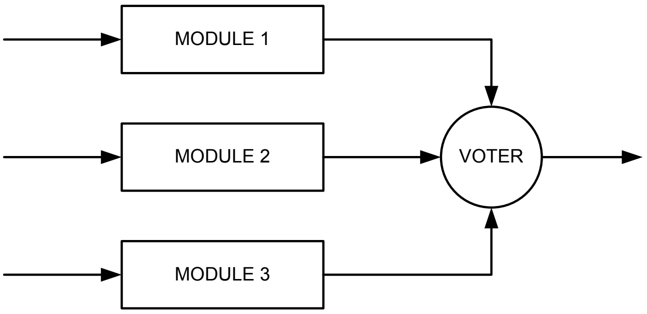


Figure 3.3: Triple Modular Redundancy.

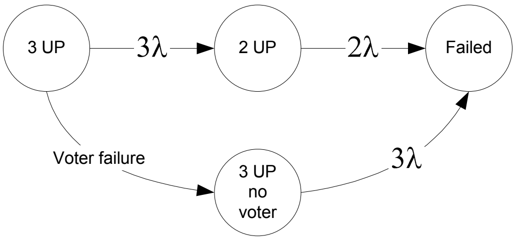


Figure 3.4: State transition diagram for the Triple Modular Redundancy.

a single point of failure and it must be very reliable and if necessary it is duplicated [82, 54]. The N-Modular Redundancy (NMR) is the generalization of the TMR, which tolerates M - 1 failures ($M < N$) out of N components. Static fault tolerance is applied for high reliability applications for which the service is non-interruptible and repair is not possible (e.g., space missions).

An example of state transition diagram for the TMR is shown in Figure 3.4. The failure rate of each identical module is assumed constant and equal to λ. The failure of one module is detected by the voter and is tolerated. The failure of the voter leaves the system uncovered so that in case of failure of one of the three modules the system fails.

3.4.2 Dynamic Fault Tolerance

Dynamic (active) fault tolerance allows for system reconfiguration after the failure. The reconfiguration process consists of detection, localization, and insulation of the fault, followed by recovery of the system to the healthy state. It may cause a downtime period during which the service is interrupted. The basic architecture is the duplication with comparison: the outputs of two parallel systems are compared and if they disagree a fault detection routine is launched after which the faulty module is recovered or replaced with a new one. **Stand-by spare** parts are used to this end. They distinguish into hot and cold stand-by spares depending if the spare unit is powered or not. Hot stand-by spares optimize the reconfiguration downtime because they are already switched on, whereas they consume energy and may fail. Cold stand-by spares have a longer reconfiguration downtime because they are switched off, whereas they do not consume energy and do not fail. The spare activation plays a crucial role like the voter for the static fault tolerant architecture [74]. Dynamic redundancy is applied for long life high availability applications (e.g., satellite communications) for which a short downtime for failure recovery is acceptable.

An example of state transition diagram for a TMR with one spare module is shown in Figure 3.5. If one module fails, this is discovered by the voter, and the spare activation starts with a probability of success C. Once the spare is successfully activated the TMR is recovered at full redundancy (state 2+1 UP), on the contrary the system continues with 2 modules (state 2 UP). The failure of the voter leaves the system uncovered (state 3 UP no voter), so that the failure of one of the three modules leads the system to the state failed[10].

3.4.3 Hybrid Fault Tolerance

Hybrid fault tolerance combines together static and dynamic fault tolerance [74]. The dynamic reconfiguration task can be applied to a static NMR scheme. The spare unit replaces the faulty module and the system is recovered to full redundancy. If this is not possible because there are no more

[10] If the voter were a single point of failure its failure would lead straightway to the state failed.

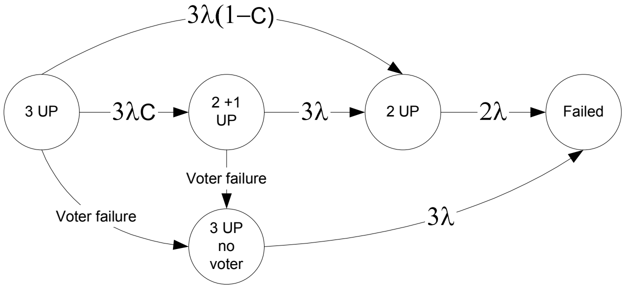


Figure 3.5: State transition diagram of a Triple Modular Redundancy with one stand-by spare.

spares available, then the pool of survived components is reconfigured after the failure in order that a majority voting still exists. For example, the 3 out of 5 architecture becomes a 2 out of 3 by removing the faulty module plus another one. The hybrid fault tolerance gives the best results but it is more complex and costly[11]. The general hybrid NMR is shown in Figure 3.6.

3.4.4 The Role of Maintenance

Maintenance is the main instrument to keep the system in a healthy status and preventing the accumulation of faults caused by aging-wearing phenomena [44]. A maintenance program will consist of preventive scheduled maintenance policies when it is performed periodically at scheduled instants, and corrective maintenance when it is performed on demand, at the occurrence of a failure [2, 84]. The preventive maintenance further specializes into wearing and aging policies. The wearing policies survey the status of the component, which is replaced if it is worn above an acceptable threshold. The aging policies replace the component at fixed intervals, called age limits, which does not depend on the status of the component. The choice of the

[11] The basic principles of software fault tolerant architectures are derived from the presented hardware architectures, like the multi-versions software with voting, though the algorithms for reconfigurations and fault recovery are more sophisticated in order to apply roll-back and roll-forward procedures [82, 74].

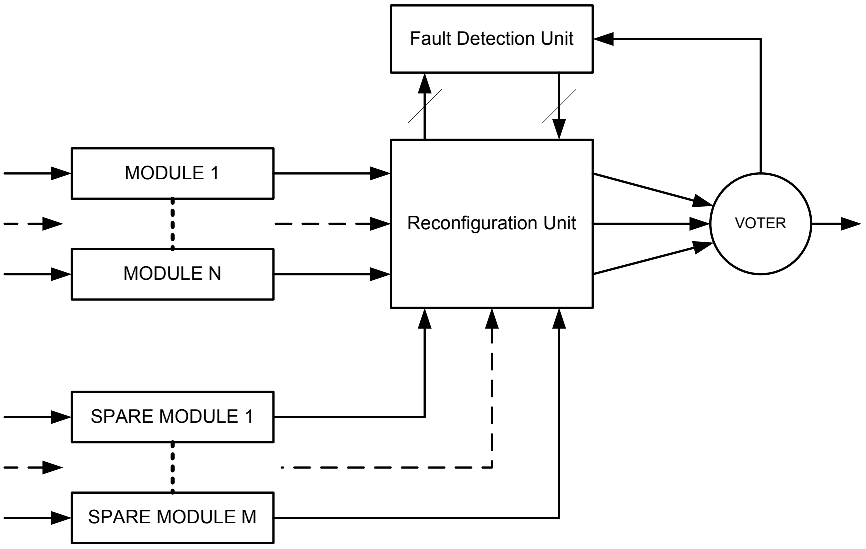


Figure 3.6: Hybrid N-Modular Redundancy with M spares.

wearing threshold, the age limit of a component and the inspection intervals is traded-off with the costs for maintenance [26].

As an example, the TMR can be periodically inspected and the faulty item repaired and restored in the pool of three components. In the model of figures 3.4 and 3.5, a transition is added from every state to the initial one. In case the dependability attribute is the reliability, then the failed state is absorbing and cannot be recovered.

3.4.5 Comparison of Fault Tolerant Architectures

Some of the presented fault tolerant architectures are compared considering their MTTF and hazard function. They are: an architecture with two modules in parallel, a TMR and a TMR with spare. The failure rate of each module λ and the failure rate λ_V of the voter are assumed to be constant. For the TMR with stand-by spare, the probability of successful spare activation is modeled with the constant C, $0 < C \leq 1$.

The dual parallel architecture has a reliability $R_{2P} = 2R - R^2$, where $R = e^{-\lambda t}$, for a $MTTF_{2P} = \frac{3}{2\lambda}$. For the other architectures the reliability can be calculated by solving the respective Markov chain, shown in Figure 3.4 and Figure 3.5 respectively, or using the analytic expression. To obtain this result are used formulas that will be presented in the next Chapter 5. The reliability of the TMR is calculated using equation (5.8):

$$R_{TMR1}(t) = \alpha \int_0^t Hyp[3\lambda + \lambda_V, 2\lambda] d\tau + \beta \int_0^t Hyp[3\lambda + \lambda_V, 3\lambda] d\tau \tag{3.6}$$

where $\alpha = (\frac{3\lambda}{3\lambda+\lambda_V})$, $\beta = (\frac{\lambda_V}{3\lambda+\lambda_V})$ and Hyp is the hyperexponential distribution[12]. The MTTF is calculated using equation (5.9):

$$MTTF_{TMR1} = \alpha(\frac{1}{3\lambda + \lambda_V} + \frac{1}{2\lambda}) + \beta(\frac{1}{3\lambda + \lambda_V} + \frac{1}{3\lambda}) \tag{3.7}$$

The maximum is obtained for $\lambda_V = 0$, for which $MTTF = \frac{5}{6\lambda}$. The reliability of the TMR with spare is calculated using equation (5.8):

$$\begin{aligned} R_{TMR2}(t) = R_{TRM1} + C(\alpha^2 \int_0^t Hyp[3\lambda + \lambda_V, 3\lambda + \lambda_V, 2\lambda] - \\ \alpha \int_0^t Hyp[3\lambda + \lambda_V, 2\lambda] + \alpha\beta \int_0^t Hyp[3\lambda + \lambda_V, 3\lambda + \lambda_V, 3\lambda]) \end{aligned} \tag{3.8}$$

from which the MTTF results:

$$\begin{aligned} MTTF_{TMR2} = MTTR_{TMR1} + \\ C(\alpha^2(\tfrac{2}{3\lambda+\lambda_V} + \tfrac{1}{2\lambda}) - \alpha(\tfrac{1}{3\lambda+\lambda_V} + \tfrac{1}{2\lambda}) + \alpha\beta(\tfrac{2}{3\lambda+\lambda_V} + \tfrac{1}{3\lambda})) \end{aligned} \tag{3.9}$$

The equations above attest, as is logical, that reliability and the MTTF of the TMR with spare is always bigger than the reliability and the MTTF of the TMR, whatever the parameter settings. The maximum is obtained for C = 1 and $\lambda_V = 0$, for which $MTTF = \frac{5}{3\lambda}$.

The three architectures are compared with respect to the resulting MTTF. The MTTF of the TMR with spare, assuming a never faulty spare activation, is the longest one followed by the dual parallel redundancy and the TMR.

[12] The hyperexponential distribution is obtained by the convolution of the density functions characterizing each state transition in the chain. This definition holds for homogeneous continuous time Markov chain, see section 5.2.

The MTTF of the TMR is even shorter than the MTTF of a single module, which is $1/\lambda$. This result is well known in literature and demonstrates mathematically that the TMR is not suited for long missions.

The three architectures are compared with respect to the hazard functions defined in equation (3.1), for $\lambda = 0.0001$, $\lambda_V = 0$ and $C = 1$. Results are shown in Figure 3.7. At the early period (< 4000 h) the hazard function (b) of the TMR with spare is below all other curves, which confirms the efficacy of this architecture already attested by the longest MTTF. In this sense, the crossing point around 3000 hours with (d) is not that significant. The hazard function (a) of the simple TMR is always larger than the dual parallel architecture (d) and of the TMR with spare (b). Nevertheless, the TMR hazard function stays below the 0.0001 failure rate of the single module (c) up to 8.000 hours[13]. Once more this justifies the utilization of this architecture for missions of medium length.

The effect of maintenance, either corrective or scheduled at fixed intervals, may further ameliorate the behavior of the analysed examples. Results are sensitive to the voter failure rate and the spare activation, which can degrade the performance of the TMR architectures. This is not only a mathematical evidence. For those systems demanded to be safe and reliable, a simple architecture is often preferred. This is because the sensitivity (i.e. robustness) of the result with respect to the other design facilities represents a factor of uncertainty on the calculated dependability attribute, which turns to be a design trade-off. For this reason, a simple parallel redundancy is preferred to more complex architectures.

[13] More properly, the crossing point is the time where the two architecture have the same reliability, which is related to the integral of hazard function. This is expected to be larger than the crossing point shown in Figure 3.7

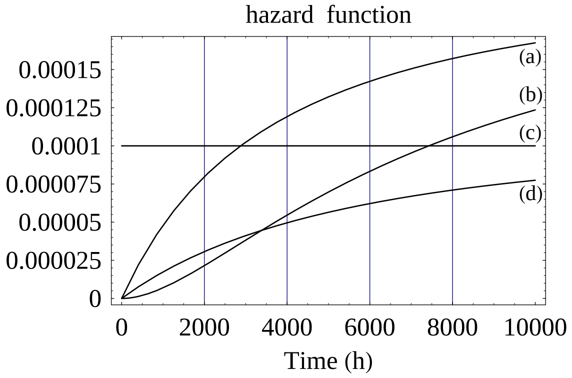


Figure 3.7: Hazard functions of the TMR (a), the TMR with spare (b), a single module (c) and two modules in parallel (d).

3.5 Design for Safety

A system is defined safety critical if it delivers at least one safety critical service [63]. A service is called safety critical if its failure may have serious consequences such as loss of human life and severe injuries (i.e. life-critical), large-scale environmental damage (i.e. environmental-critical) or economical penalties (i.e. costs-critical) [46]. These definitions apply to many fields like military, industry, transports, emergency communications, medicine, nuclear and chemical plant and space missions[14].

A safe design is based on some preliminary recommendations. First of all, the system must be testable. As this is easier for hardware than for software, the safety critical systems are preferably hardware. Secondly, no single points of failure must exist so that back-up parallel systems, possibly using different technology, are recommended [46]. Thirdly, as the man-machine interface is a major concern for safety, it must be assured that the risk to harm the system by non-intentional human errors is reduced to a minimum. As a result of the safe design, a system is classified inherently safe, fail safe (i.e.

[14] Safety only deals with non-malicious failures, i.e. those originated either by the system failure process or by flaws in the design and the production. A hazardous state can be reached intentionally but in this case it concerns security and not safety.

Category	**Likelihood**	**Frequency/year**
Frequent	Very likely to occur event	> 1
Probable	Likely to occur event	0.1 - 1
Occasional	Possible and expected to occur event	0.01 - 0.1
Remote	Possible but unexpected to occur event	0.001 - 0.01
Improbable	Unlikely to occur	0.0001 - 0.001
Negligible	Extremely unlikely to occur	< 0.0001

Table 3.1: Frequency categories.

the mission shutdown), fail operational (i.e. emergency back-up units) and fail soft (graceful degradation). Inherently safe systems deal with intrinsic safety, while the others deal with engineered safety [63].

Design for safety deals with the reduction of the frequency of the hazard and the control of the consequences, see Tables 3.1 and 3.2. The product of the hazard frequency by the consequences is the risk, see Table 3.3. The risk is quantified by **risk analysis** [5], the result of which returns a maximum failure rate or a failure probability that the system must satisfy in order to be certified safe[15]. The IEC 61508 standard ranks safety in four categories, the Safety Integrity Levels (**SIL**), from SIL1 to SIL4, the strictest one [42, 5] to which corresponds four classes of risk from I to IV, see Table 3.4. The class of risk I is the highest and corresponds to an intollerable risk. The classes of risk II corresponds to a undesiderable risk that is tolerable only if its reduction is impracticable or excessively costly with respect to the gained improvement. The class III corresponds to a tolerable risk obtained by a reasonable risk reduction[16]. Class IV is a negligible risk.

The systems that contribute to achieve such a SIL are defined in literature as Safety Related Systems (**SRS**). Two types of SRS exist: SRS for continuously controlled systems and SR protection systems. Both of them drive the system into a safe state from where the system cannot arbitrarily

[15] As an example, high safety critical applications require a probability of failing unsafe below 10^{-9} for 10 hours mission [74].

[16] Qualitative guidelines to judge a risk of II or III type are given by the As Low As Reasonably Predictable (ALARP) principle [5].

Category	**Gravity**		**Damage**	
	Gravity	**N. fatalities**	**Loss(CHF)**	**Downtime**
Catastrophic	Multiple fatalities	> 1	$>$ 100 MCHF	$>$ 3 months
Major	Single fatalities	1	1-100 MCHF	1 week - 3 months
Severe	Non fatal injuries	0.1	0.01 - 1 MCHF	4 hours - 1 week
Minor	Minor injuries	0.01	0 - 10 KCHF	$<$ 4 hours

Table 3.2: Consequence categories.

Frequency	**Consequences**			
	Catastrophic	Critical	Marginal	Negligible
Frequent	I	I	I	II
Probable	I	I	II	III
Occasional	I	II	III	III
Remote	II	III	III	IV
Improbable	III	III	IV	IV
Incredible	IV	IV	IV	IV

Table 3.3: Frequency x consequences = class of risk.

SIL	**SR Control systems**	**SR protection systems**
	Failure rate/h	Prob. of failure on demand
4	$[10^{-9}, 10^{-8}]$	$[10^{-5}, 10^{-4}]$
3	$[10^{-8}, 10^{-7}]$	$[10^{-4}, 10^{-3}]$
2	$[10^{-7}, 10^{-6}]$	$[10^{-3}, 10^{-2}]$
1	$[10^{-6}, 10^{-5}]$	$[10^{-2}, 10^{-1}]$

Table 3.4: Safety Integrity Levels.

move [63]. The first type of SRS maintains the system in a safe state by activating a back-up unit that continues to perform the vital functions as long as it is necessary (e.g., the fly by wire control systems of an airplane). The second type of SRS works on demand and leads the system into a safe state by a mission shutdown as soon as a critical event has been detected (e.g. the protection system of a nuclear power plant) [45, 63]. The SIL for the two categories of SRS are shown in the two columns of Table 3.4.

These days, the SRS are PLC based architectures. A rich literature in standards and manuals exists that address the subject like the IEEE guide for nuclear power station protection systems [45], the IAEA safety guide for protection systems [41] and the IEC standard for safety related systems [42].

3.5.1 Failsafe Protection Systems

The architectures for safety critical applications are all characterized by a protection system that surveys those events that may potentially impair the system safety. Parts of these event are detected by monitoring a physical process, for example an over temperature or an over pressure. Another part concern the monitoring of the status of the components. The monitored events are the **initiating events** of the protective action [41]. Once an initiating event has been detected, the **protective action** is transferred to a **safety actuation system** that moves the system into a failsafe state. For a protection system of a nuclear power plant the protective action generates the mission abort and drops the graphite rods into the core of the reactor to stop the reaction. For the LHC, the protective action generates the mission abort by removing the beams from the rings, as explained in section 1.3.

The temporary unavailability of a protection system does not directly harm safety. The system still functions, though subject to higher risks [5]. Nevertheless, every protection system should have self-monitoring facilities implemented that in case of detected internal failure, trigger a safe shutdown avoiding the possibility the protected system might function at an uncovered risk. A shutdown because of safety reasons is called a false shutdown, which

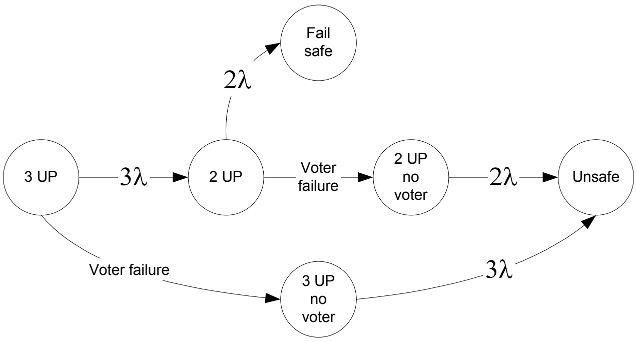


Figure 3.8: The state transition diagram for a failsafe Triple Modular Redundancy.

is a false trip in the nuclear power plants or a false beam dump in the LHC. The safety is a trade-off with respect to other dependability attributes, like availability and reliability, which usually gets penalized by the failsafe strategy of aborting the operation [22, 74]. As an example of a failsafe architecture, the TMR can be transformed failsafe if the disagreement among the outputs, detected by the voter, results in a failsafe shutdown. This is modeled in the state transition diagram of Figure 3.8 for the TMR of Figure 3.4. Similar modifications apply to the model of Figure 3.5.

Chapter 4

Probability Models

This chapter introduces to the theory of renewal stochastic processes, the Markov processes and the Markov regenerative processes. Foundaments on probability theory are omitted as they can be found in many textbooks [39, 80, 87].

4.1 Generalities on Stochastic Processes

A stochastic process is a family of random variables $X(t)$, called states of the process, defined in the probability space X, with t $\in$ T, where T is the time index. The stochastic processes are classified into 4 categories depending on T and the state space X, which can be discrete or continuous [87]. For both T and X being continuous, a realization of the process is a function $x(t)$ for $t \in [0, T)$. For X being discrete, a realization of the process is a sequence of n ordered outcomes $\{x_i, t_i\}$, i = 1,2…n, and the stochastic process is called a chain.

Definition 4.1. A stochastic process is fully characterized with its n-order probability distribution $P : X \times T \to [0, 1]$, which is defined as:

$$0 \leq F[\bar{x}, \bar{t}] = P[X(t_1) \leq x_1, \ldots, X(t_n) \leq x_n] \leq 1 \tag{4.1}$$

where $\bar{x} = (x_1, \ldots, x_n)$ are n samples of $X(t)$ at the instant $\bar{t} = (t_1, \ldots, t_n)$.

The process is called strict stationary if equation (4.1) is invariant to the time shift $t+\tau$ for all n and $\tau > 0$. It is called wide stationary if the average $E[X(t)]$ is constant and the autocorrelation function is invariant to the time shift that is $R(t+\tau,\tau) = E[X(t+\tau)X(\tau)] = R(t)$ [87]. The process is called ergodic if equation (4.1) is constant or asymptotically converges to a constant steady state distribution.

The process is driven by dependencies existing within the states to reach the next and all or part of the visited states, i.e. the history of the process. A further classification is possible considering this feature:

Definition 4.2. Stochastic processes are independent if the n-order distribution is deduced from the first order distribution.

$$P[\bar{x},\bar{t}] = P[X(t_1) \leq x_1, \ldots, X(t_n) \leq x_n] = \prod_{i=1\ldots n} P[X(t_i) \leq x_i] \tag{4.2}$$

Definition 4.3. Stochastic processes are Markov if the n-order distribution is deduced from the first order conditional distribution.

$$\begin{aligned} P[\bar{x},\bar{t}] &= P[X(t_1) \leq x_1, \ldots, X(t_n) \leq x_n] = \\ &= \prod_{i=2\ldots n} P[X(t_i) \leq x_i | X(t_{i-1}) \leq x_{i-1}] P[X(t_1) \leq x_1] \end{aligned} \tag{4.3}$$

The equations (4.1), (4.2) and (4.3) hold if X is a partial ordered set, which is not necessarily true for X being discrete. In this case, (4.2) and (4.3) have to be rewritten as follows:

$$P[\bar{x},\bar{t}] = \prod_{i=1\ldots n} P[X(t_i) = x_i] \tag{4.4}$$

$$P[\bar{x},\bar{t}] = \prod_{i=2\ldots n} P[X(t_i) = x_i | X(t_{i-1}) = x_{i-1}] P[X(t_1) = x_1] \tag{4.5}$$

Equations (4.4) and (4.5) describe the probability of a single realization in X, T. The time that the process spends in each state is called sojourn time.

4.2 Renewal Processes

Renewal processes deal with random events that occur as statistically identical replicas of themselves, independently from the history of the process. The formal definition follows:

Definition 4.4. A renewal process is a process of independent identically distributed random variables $X(t_k)$ [87] for which the equation (4.2) describes the probability of a single realization.

Some examples of renewal processes are given assuming X to be discrete. A **renewal sequence** $\{X_i, T_i, i \geq 0\}$ is a renewal process for discrete T and generally distributed sojourn times. For X equal to {0,1}, the sequence $\{X_k, k = 1, 2, 3 \ldots\}$ is called a Bernoulli process. The sum $S_n = X_1 + \ldots + X_n$ of n consecutive outcomes of a Bernoulli process is a random variable in {0,1, . . . , n}, which describes a **discrete time renewal counting process** called also Binomial process. An example is shown in Figure 4.1, where the probability p is the probability of leaving the state and 1-p is the probability of remaining in the state.

The sum $S_n = X_1 + \ldots + X_n$ in [0, t) of the outcomes of a renewal process for continuous T and X = {0,1} is the random variable $N(t) = n$ for $t > 0$ which describes a **continuous time renewal counting process**. An example is shown in Figure 4.1. where rate λ is the instantaneous conditional probability of leaving the state. The average number of renewals in [0, t) is the renewal function $M(t) = E\{N(t)\}$ and is calculated with the fundamental renewal equation:

$$M(t) = F(t) + \int_0^t M(t-x)dF(x) \tag{4.6}$$

where $F(t)$ is the probability that $X = 1$ at time t. The time derivative of (4.6) is the renewal density $m(t)$, described by the renewal equation:

$$m(t) = \frac{dM(t)}{dt} = f(t) + \int_0^t m(t-x)f(x)dx = f(t) + m(t) \otimes f(t) \tag{4.7}$$

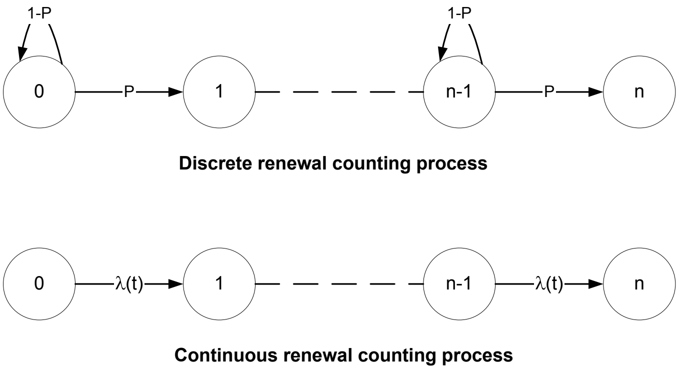

1-P
0
P
1
1-P
n-1
P
n
Discrete renewal counting process
0
λ(t)
1
n-1
λ(t)
n
Continuous renewal counting process

Figure 4.1: Renewal counting processes.

where $f(t)$ is the time derivative of $F(t)$ and the symbol $\otimes$ indicates the convolution operator[1].

The renewal continuous time counting process is called Homogeneous Poisson Processes (**HPP**) if the sojourn times are exponentially distributed with constant rate λ. In this case, the variable $N(t)$ is Poisson distributed with renewal density λ [87]. The Non-Homogeneous Poisson Process (**NHPP**) generalizes the counting processes allowing for general distributed sojourn times though it loses the renewal property [39]. For a renewal counting process the time is reset at the occurrence of the renewal event. This is not true for the NHPP for which the sojourn time between two consecutive events depends on the absolute time t of the process, which is never reset.

Other renewal processes can be obtained if the set X is taken different from {0,1}. For X = $\{1, 2, 3, \ldots\}$ the sum S is known as random walk. If negative values for X are also allowed, then the process describes a Brownian motion[2] [80].

The alternating renewal processes are a generalization of the renewal processes, where two or more variables with their distributions alternate each other. For example, the failure of a component and its repair is an alternat-

[1] Both equations (4.6) and (4.7) can be transformed in the Laplace domain [39, 87].

[2] For T being a real number and x(t) Gaussian distributed, the Brownian motion is known as white noise.

ing renewal process. On the contrary, the superposition of renewal processes is not a renewal process, as this class is not closed to this operator with the only exception of the superposition of Poisson processes that remains a Poisson process [39, 87].

4.3 Markov Processes

Markov processes deal with discrete event systems, the dynamic of which is governed by an elementary cause-effect mechanism among neighbour states. The memory-less property states that the future evolution of the process depends on the last reached state and, in general, on the time t when this state has been reached.

Markov chains are Markov processes with continuous T and discrete X, either finite or countable infinite. The equation (4.5) describes a Non-Homogeneous Continuous Time Markov Chain (**NHCTMC**). If the conditional probability depends only on the time that the process spends in each state and not on the instant when they are reached, then an Homogeneous Continuous Time Markov Chain (**HCTMC**) is obtained. The sojourn times of the HCTMC are exponentially distributed and the equation (4.5) can be rewritten as:

$$P[\bar{x}, \bar{t}] = \prod_{k=2 \ldots n} P[X(\tau_k) = x_k | X(0) = x_{k-1}] P[X(t_1) = x_1] \tag{4.8}$$

where $\tau_k = t_k - t_{k-1}$. The (4.8) describes the probability of a realization of the process, namely the sequence of states x_k visited at the instants t_k. It can be used to calculate either the probability that the final state x_k is reached from x_1 at time $t = t_k$ for the given sequence of states, whatever the timing, or the probability the state x_k is reached at $t = t_k$ from x_1 whatever the sequence of states and the timing.

For a finite state space X of size N, the total probability to be in the state x_k at time t, given the initial state x_1 at time t_1, is described by the

Chapman-Kolgomorov equation [87]:

$$P[X(t) = x_k] = \sum_{k \in X} p_{ik}(t, t_1) P[X(t_1) = x_1] = P[X(t_1) = x_1] V_k(t, t_1) \quad (4.9)$$

where $V(t, t_1)$, $t > t_1$ is the $N \times N$ matrix of the conditional probabilities, with $V(t, t) = I$, the identity matrix. The probability distribution in X at time t is also calculated by the **Kolgomorov** equations [87, 80], which derives from (4.9):

$$\frac{dP(t)}{dt} = P(t)Q(t) \quad (4.10)$$

where $\sum_{k \in X} P_k(t) = 1$ is the condition of normalization[3]. The equation (4.10) calculates the **transient solution** of the Markov chain for the initial distribution $P(0) = P_0$. $Q(t)$ is the matrix of the transition rates[4] that is defined as:

$$Q(t) = \lim_{\Delta t \to 0} \frac{V(t + \Delta t, t) - I}{\Delta t} \quad (4.11)$$

The following properties for the matrix $Q(t)$ hold:

(i) $q_{ij}(t) \geq 0$, $\forall x_i \neq x_j$,

(ii) $q_{ii}(t) \leq 0$, $\forall x_i$,

(iii) $\sum_{x_i \in X} q_{ij}(t) = 0$,$\forall x_i$.

The items i) and ii) state that the input rates are positive for all states and the output rate is negative. The item iii) is the balance equation of the input-output rates that holds only for closed Markov chains, i.e. those ones that have neither sources nor sinks. For the HCTMC, the matrix Q is constant and $V(t) = \exp(Qt)$[87].

The **steady state** solution of (4.10) is the state probability distribution $P(t)$ for $t \to \infty$. For the HCTMC, it is obtained by the solution of the system of linear equations $PQ = 0$, which does not depend on the initial condition

[3] The normalization holds only for closed Markov chains.

[4] The Q accounts for the instantaneous conditional probabilities and respects the Markov property, while $V_{jk}(t)$ is the cumulative conditional probability that accounts for all the paths leading to the state x_k from the state x_j in time interval t.

and for this reason is said ergodic. In general, for the NHCTMC, the steady state solution depends on the initial condition.

A Markov Chains can be studied with respect to the **structural properties** of its matrix Q. This analysis makes it possible to classify the states with respect to their reachability (i.e. looking at the Markov chain like an oriented graph) and the expected dynamics. Some definitions follow:

Definition 4.5. A state x_k is reachable from the state x_j if there exists at least one path, i.e. a sequence of oriented arc-transitions, leading to it.

On the basis of this definition, the chain is called irreducible (or cyclic) if every state is reachable from every other state, otherwise it is called reducible. The transition matrix of a Markov chain is usually a sparse blocks matrix. In the special case where each state is reachable from the others in only one step, Q is a dense matrix.

Definition 4.6. A state is called transient if there is a positive probability the system will not return in it, and it is called recurrent if the probability to return in that state is unitary. A recurrent state is called absorbing if its output rate is null.

The transient analysis usually applies to Markov chains with X partitioned in absorbing and transients states, for which the analysis at the steady state is less interesting. On the contrary the steady-state analysis applies to irreducible Markov chains, where each state is recurrent, for which the transient solution is less interesting [80, 87, 18].

Markov chains with infinite countable states underlie the same mathematics of the finite Markov chains but they are harder to solve. Solutions exists if a repetitive structure exists, as it is the case of elementary queuing and failure-repair models[5]

[5] A birth-death process has as state variable $X = n, (n \geq 0)$ that represents the balance births - deaths at time t. The steady state exists if the chain is stable, namely the birth/death ratio is less than 1 [87].

The class of Markov process is large enough to include some of the renewal processes presented in section 4.2. For example, a Binomial process is a pure death discrete time Markov chain (**DTMC**), a Non Homogeneous Poisson Process is a **NHCTMC** and a random walk is a Markov chain.

4.4 Markov Regenerative Processes

Markov regenerative processes deal with processes that regenerate at determinate instants according to an embedded Markov-like dynamic [59]. The formal definition is:

Definition 4.7. A stochastic process $X(t), t \geq 0$ is defined a Markov Regenerative Process **MRGP** if there exists a Markov Renewal Sequence **MRS** $\{Y_i, T_i, i \geq 0\}$ of random variables Y embedded in the process, where $Y \in S \subseteq X$, such that all conditional finite dimensional distributions of $\{X(T_i + t), t \geq 0\}$ given $\{X(u), 0 \leq u \leq T_i, Y_i = m\}$ are the same as those of $\{X(t), t \geq 0\}$ given $Y_0 = m$.

The process is regenerated at given epochs, called the regeneration points, which are the outcomes of the MRS. In general, more than one regeneration event can be enabled at time t. Within one epoch, the variable $X(t)$ changes according to the subordinated process, which is not necessarily Markov. At the triggering of the regeneration event, the reached state is frozen and the process restarts in the new epoch with regenerated dynamic[6].

The **transient solution** of the **MRGP** [59] gives the probability distribution in $X = \{1, 2, 3 \ldots N\}$ at time t. The matrix V(t) of equation (4.9) is the solution of the **generalized Markov renewal equation**:

$$V(t) = L(t) + \int_0^t V(t-u)dK(u) = L(t) + V(t) \odot K(t) \tag{4.12}$$

[6] This does not imply that the state probability distribution or any cumulated statistics go back to the initial value. Only the instantaneous conditional probabilities do.

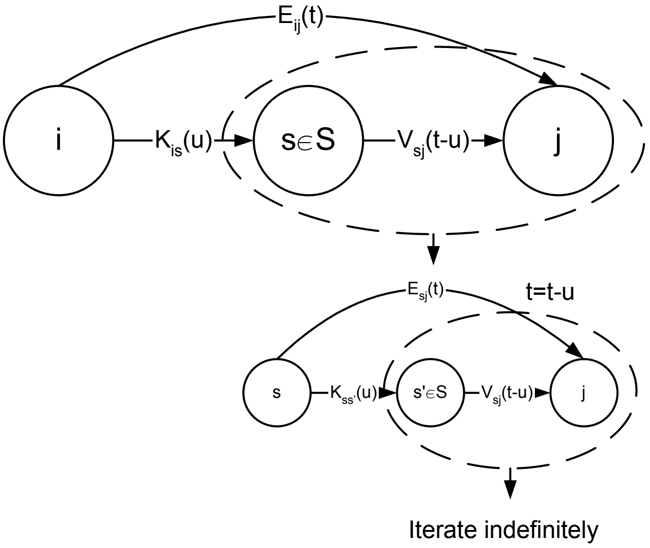


Figure 4.2: A realization of a Markov regenerative process.

for the initial conditions $V(0) = I$. The operator $\odot$ is the Laplace-Stieltjes convolution. The (4.12) is a set of coupled Volterra integral equations of the second kind, which can also be transformed in the Laplace domain [59].

The $N \times N$ matrix $K(t)$ is the **global kernel** of the MRGP and governs the occurrence of the regenerations epochs. It is defined as:

$$K_{ij}(t) = P[Y_1 = j, T_1 \leq t | Y_0 = i] \tag{4.13}$$

The $N \times N$ matrix $L(t)$ is the **local kernel** of the MRGP and governs the conditional probability distribution of subordinated process in one epoch. It is defined as:

$$L_{ij}(t) = P[X(t) = j, T_1 > t | Y_0 = i] \tag{4.14}$$

An intuitive representation of equation (4.12) is given in Figure 4.2. The state $X = j$ is reached from the state $X = i$ from two paths: one is direct, within one epoch, while the second passes through one or more regeneration epochs. For simplicity, only the initial state i, the final state j and the regeneration states s $\in S$ are shown, though the system is free to move into others states in X, according to the local kernel $L(t)$.

The **steady state solution** of an MRGP exists if the MRS is a-periodic and irreducible. The steady state of the MRS is calculated by solving the system of equations $Y = Y \times \lim_{t\to\infty} K(t)$, normalized in S, that is $\sum_{s\in S} Y_s = 1$ and calculating $A = \int_0^\infty L(t)dt$. The steady state of the MRGP for the state j is :

$$\pi_j = \frac{\sum_{s\in S} Y_s A_{sj}}{\sum_{s\in S} Y_s \sum_{r\in S} A_{jr}} \tag{4.15}$$

which is the sum of the product of the fraction of time spent in each state of the MRS multiplied by the time spent in j, divided by the total time spent in each visit of j [59].

4.4.1 Examples of Markov Regenerative Processes

In its general formulation, the MRGP is very hard to solve except for some special conditions that, if verified, help to keep the problem mathematically treatable. They are: 1) a simple structure of the MRS (e.g. a-cyclic, non recurrent), 2) the subordinated process that is Markov and 3) the independence between the MRS and the subordinated process. Some examples from literature follow.

The **Semi-Markov** processes (**SMP**) are MRGP for which each state transition is a regeneration point. As a consequence, the matrix $L(t)$ is diagonal [59].

$$L_{ij}(t) = [1 - \sum_{j\in S\equiv I} K_{ij}(t)]\delta_{ij} \tag{4.16}$$

where $\delta_{ij} = 1$ if $i = j$, otherwise it is zero. An SMP can be deduced from a general MRGP. Given an embedded renewal Markov sequence $\{Y_i, T_i\}$, then the continuous process $X(t), t \geq 0$ such that $X(t) = Y_{N(t)}$, $N(t) = \sup(n \geq 0, 0 \leq T_n \leq t)$, is an SMP. If the Markov renewal process is a Semi-Markov chain $\{1,\ 2,\ \ldots \mathrm{M}\}$, with only one regeneration event active at time, then $V(t)$ becomes:

$$V(t) = L(t) + K_1(t) \odot (L(t) + K_2(t) \odot (\ldots) \tag{4.17}$$

In literature, there exist cases of MRGP that also applies to the solution of special classes of **Petri nets**. If the subordinated process is HCTCM then the overall MRGP is demonstrated to underlie the class of the Markov Regenerative Stochastic Petri nets [21]. In case the MRS is a sequence of deterministically distributed variables, enabled one at time and never preempted, then the MRGP underlies the class of the Deterministic Stochastic Petri nets and the solution of (4.17) reduces to a series of matrix multiplications [20, 69]:

$$V(t) = \left(\prod_{h=1}^{j-1} \exp[Q\tau_h]\Delta_h \right) L(t - \sum_{h=1 \ldots j-1} \tau_h) \tag{4.18}$$

where τ_h are the deterministic sojourn times in the state h of the MRS, and Δ is the branching probabilities matrix that accounts for the instantaneous state mapping between two consecutive regeneration epochs.

4.5 Markov Reward Processes

The Markov reward processes extend the modeling capability of all previously illustrated models thanks to the possibility of defining reward functions of the states of the system. A more formal definition follows:

Definition 4.8. A reward w_x is a rate assigned to each state X of the stochastic process (rate-type reward) or to the transitions (impulse-type rewards).

Definition 4.9. The cumulative reward function $W(t)$ is a function of the rewards accumulated over t in the state space X.

Definition 4.10. A Markov Reward Process **MRP** is a Markov process for which reward rates are associated either with states (rate-type rewards) or with transitions (impulse-type rewards) [88].

Only rate-type rewards are considered here. The Markov reward processes have constant reward rates assigned to each state of the chain, so that the cumulative reward function $W(t)$ is the sum of all rewards cumulated in each state of X. The formulas of the expected instantaneous reward rate and the expected cumulative reward are the following [59]:

$$E[w(t)] = \sum_{i \in I} w_i P_i(t) \tag{4.19}$$

$$E[W(t)] = \sum_{i \in I} w_i \int_0^t P_i(\tau) d\tau \tag{4.20}$$

The integral of $P(t)$ is the cumulative state probability vector of the MRP and denotes the expected total time spent by the process in a state during the interval [0, t). To compute this quantity, it is necessary to solve the following equation, derived from (4.10):

$$\frac{d\Pi(t)}{dt} = \Pi(t)Q + P(0) \tag{4.21}$$

where $\Pi(t) = \int_0^t P(x)dx$.

The reward rate at the steady state are obtained by multiplying the vector w by the steady-state distribution [88]. Reward processes can be built on Semi-Markov [24] and Markov regenerative processes [59] as well. For the Markov regenerative reward process the variable $w(t)$ is evaluated through different epochs, that is:

$$E[w(t)] = [V(t)w^T]P(0) \tag{4.22}$$

where $V(t)$ is calculated with (4.12). A generalization of the MRP is the non homogeneous MRP (**NHMRP**), for which the reward rate may be a function of time t.

A reward processes can be defined on the models presented in the previous sections of this chapter. As an example, the compound **HPP** is a class of reward processes built on the Poisson processes, where the cumulative rewards W_1, W_2, etc. are independent random variable with identical probability distribution associated to the states 1, 2, etc.... The reward function is $W(t) = \sum_{i=1...N(t)} W_i$. The Wald's theorem provides the average and the variance of $W(t)$ equal to $E\{W\}\lambda t$ and $\lambda t(E\{W\}^2 + Var\{W\}^2)$ respectively

[39]. The renewal reward processes are a generalization of the compound HPP.

Reward process are suited for performability analysis, with the reward function representing the performance index of the system. They can also define dependability attributes. For example, assuming a reward rate equal to 0 for the states where the system has failed and a reward rate equal to 1 for the states where the system is functioning defines a reward function that corresponds to the system reliability.

4.6 The Solution of a Markov Process

The analysis of a stochastic process passes through the choice of the solution method, the characterization of the result as function of the uncertainty in the parameters and its sensitivity to small perturbations.

There exist various methods to obtain the solution of the Markov chain. A sample of these are listed in Table 4.1 with their asymptotic complexity[7] $O()$ and applicability. The transient solution of a Markov chain, with finite, discrete X and continuous T, is the expression:

$$P(t) = V(t,0)P(0) \tag{4.23}$$

where $P(t)$ is a vector of size N and $V(t,0)$ is a $N \times N$ matrix depending on $Q(t)$ through the Peano-Baker series [79]. In general, the expression is not analytically treatable, with the exception of a constant Q. In this case, the solution can be obtained either using Laplace or through spectral decomposition (i.e. looking for the eigenvalues and eigenvectors of Q) [87]. These techniques are effective for small N and well conditioned matrix. In the more general case other methods apply, which are based on the series expansion of $V(t,0)$, like the uniformization technique [80]. The solution can also be achieved by standard numerical methods (e.g. trapezoidal rule

[7] The asymptotic complexity give the upper bound of the complexity of an algorithm g(), so that $g()$ is $O(f)$, with $O(f) = \{g : N \to R | \exists c > 0, \exists n_0 \in N, \forall n > n_0 : g(n) \leq c \times f(n)\}$

Technique	**Solution**	**Complexity**	**Remarks**
Laplace transform	Analytical	$O(N^5)$	Not cyclic Markov chains
Spectral decomposition	Analytical	$O(N^4)$	Not large Markov chains
Uniformization	Numerical	$O(N^3 \log(qt))$	Stiffness problems
Num. methods	Numerical	$O(N^2 qt)$	Apply to stiff and large chains

Table 4.1: The solution techniques for Markov chains.

based) and similar that may cope with the stiffness of Q at the desired level of accuracy. A survey of the solution techniques can be found in [78] while some special cases are treated in [6] for stiff Markov chains, [81] for adaptive uniformization, [58] for semi-Markov process, [79] for NHCTMC and [35] for MRGP.

The solution also depends on the accuracy of the parameters of the matrix Q. In reliability and more in general in dependability problems modeling the system failure processes, the parameters of Q are related to the failure rates, which are average values taken from standard literature, see section 3.2. As a consequence, the obtained solution is more properly an average solution. In literature there are few studies aiming at characterizing statistically the solution on the basis of the statistics of the parameters [38]. This analysis is usually addressed by Monte Carlo simulation [95].

In absence of statistics on the parameters of Q, a sensitivity analysis may give similar and even more general information. Sensitivity studies are made by perturbing the model with respect to a parameter and measuring the difference from the original solution[8] [76]. The sensitivity analysis can be performed either for a single parameter variation or for multi-parameter variations. The sensitivity of the state probability vector $P(t)$ with respect to a single parameter λ is defined the partial derivative of the probability vector $S_P(t) = \partial P(t)/d\lambda$. By substituting the definition in equation (4.10) it is obtained:

$$\frac{dS_P(t)}{dt} = P(t)F + S_P(t)Q(t) \tag{4.24}$$

[8] This perturbation is called structured. The perturbation is called unstructured if it is due the round-off error, which affects the entire calculation.

where $S_P(0) = 0$ and $F = \partial Q(t)/d\lambda$. If Q is constant the solution is:

$$S_P(t) = P(0) \int_0^t e^{Q\tau} F e^{Q(t-\tau)} d\tau \tag{4.25}$$

The sensitivity to multi-parameter requires the definition of the structured perturbation $\delta_s(t)$ as a function of the perturbation of the transition matrix $Q_s = Q + \Delta_s$:

$$\delta_s(t) = P_s(t) - P(t) = P(0)(e^{Q_s t} - e^{Qt}) + \delta_s(0)e^{Q_s t}$$

For small variations of the parameters the following formula holds:

$$\delta_s(t) \cong \sum_i \Delta\lambda_i \frac{\partial P(t)}{d\lambda_i} = \Delta\lambda_i S_i(t) \tag{4.26}$$

which depends on the solution of (4.24). As an alternative, sensitivity bounds can be defined on the norm of the matrix $S(t)$ [76], where the bound for a single parameter variation is $\|S(t)\| \leq t \left\|\frac{\partial Q}{d\lambda}\right\|$.

Similar sensitivity formulas exist in literature for the Markov reward process [76, 38], the semi-Markov and the Markov regenerative process [60]. In addition to the techniques illustrated so far there exist other techniques whose main goal is to treat the state explosion problem. They all apply states aggregation and composition on the basis of the structural properties of the matrix Q of the Markov chain [51], defined in section 4.3. The original chain results split into smaller sub-chains, which are separately solved in the respective state space.

Chapter 5

Dependability Modeling

This chapter applies the theory of chapter 4 to the modeling of the failure mechanisms for reliability, availability and safety applications.

5.1 Modeling of Elementary Failure Mechanisms

A Markov chain with constant failure rates describes the failure process of two components A and B, see Figure 5.1, where λ_{A1} and λ_{B1} are the failure rates of A and B in the initial state, λ_{A2} is the failure rate of A after the failure of B and λ_{B2} the failure rate of B after the failure of A. The state probability distribution is given by equation (4.10):

$$\frac{d}{dt}\mathbf{p}(t) = \mathbf{p}(t)\begin{pmatrix} -\lambda_{A1}-\lambda_{B1} & \lambda_{B1} & \lambda_{A1} & 0 \\ 0 & -\lambda_{A2} & 0 & \lambda_{A2} \\ 0 & 0 & -\lambda_{B2} & \lambda_{B2} \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.1}$$

which is solved for the initial state probability vector $\mathbf{p}(0) = [1, 0, 0, 0]$. The probability to be in the failure state is also given by the formula:

$$P_{A\cap B}(t) = \alpha \int_0^t Hyp(\lambda_{A1}+\lambda_{B1}, \lambda_{B2})d\tau + \beta \int_0^t Hyp(\lambda_{A1}+\lambda_{B1}, \lambda_{A2})d\tau \tag{5.2}$$

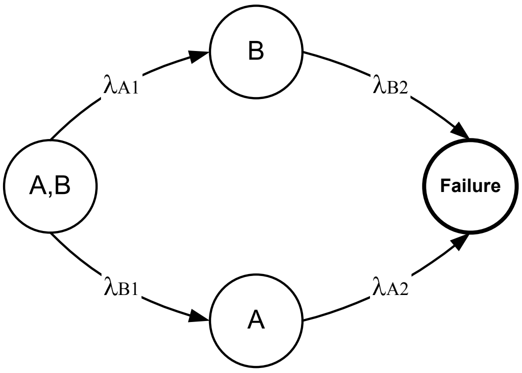


Figure 5.1: Markov chain for modeling failure dependence.

where $\alpha = \frac{\lambda_{A1}}{\lambda_{A1}+\lambda_{A2}}$, $\beta = \frac{\lambda_{B1}}{\lambda_{A1}+\lambda_{A2}}$, and $Hyp(\lambda_1, \lambda_2) = \frac{\lambda_1\lambda_2}{\lambda_1-\lambda_2}[\exp(-\lambda_1 t) - \exp(-\lambda_2 t)]$ is the hypoexponential distribution [87]. The two terms in (5.2) accounts for two separate contributions to the system failure, one through the state B, given A has failed first, and the other through the state A, if B has failed first. The Mean Time To Failure (MTTF) is:

$$MTTF = \int_0^\infty [1 - P_{A\cap B}(\tau)]d\tau = \frac{1}{\lambda_{A1} + \lambda_{B1}}(1 + \frac{\lambda_{A1}}{\lambda_{B2}} + \frac{\lambda_{B1}}{\lambda_{A2}}) \tag{5.3}$$

The presented model is general enough to define the elementary failure mechanisms occurring between two components.

Definition 5.1. Two components A and B are called statistically independent if the failure of A does not depend on the failure of B and vice versa, that is $P_{A\cap B} = P_A P_B$.

The model of Figure 5.1 for A and B being independent has $\lambda_{A1} = \lambda_{A2}$ and $\lambda_{B1} = \lambda_{B2}$. The $MTTF = 1/\lambda_A + 1/\lambda_B - 1/(\lambda_A + \lambda_B)$.

Definition 5.2. Two components A and B are called positive dependent if the failure rate of one component increases after the failure of the other component so that $P_{A\cap B} = P_{A|B}P_B > P_A P_B$, where $P_{A|B}$ is the conditional probability that A has failed given B has failed [39].

The **cascade and propagation failures** are an example of positive dependent failures, where the failure of a component accelerates the failure rate of the other. The **common mode failures** are a case limit of positive dependency where the failure of one of the two components causes (deterministically) the failure of the other component[1]. In the model of Figure 5.1 this is accounted for by assuming $\lambda_{A2} = \infty$ ($\lambda_{B2} = \infty$). Another example of positive dependency is represented by the fault detection mechanism. The failure of A is possible only if the other component B, the fault detector, has failed before. This is accounted for by assuming $\lambda_{B2} = 0$. In this case, the state B is the state for the successful detection. The probability of failure is an ordered statistics[2]:

$$P(A|T_B < T_A) = \int_0^t P_B(t) dP_A(t) \tag{5.4}$$

for which $MTTF = \frac{\lambda_B}{\lambda_A+\lambda_B}(\frac{1}{\lambda_A+\lambda_B} + \frac{1}{\lambda_A})$

Definition 5.3. Two components A and B are called negative dependent if the failure rate of one component decreases after the failure of the other component so that $P_{A\cap B} = P_{A|B}P_B < P_A P_B$.

Concurrent and mutual exclusive events are examples of negative dependency, where the failure of one components inhibits the failure of the other. In the model of 5.1 this is represented with $\lambda_{A2} = \lambda_{B2} = 0$.

5.2 Models for Reliability

Reliability models splits the state space X into transient states and absorbing states. The set of the unreliable states U consists of absorbing states. As a consequence, the Markov chain for reliability models has a singular transition matrix $Q = [Q_{n\times m}, 0_{n\times(n-m)}]^T$. The reliability is defined as the probability to be in the subset of the reliable states $X|U$ at time t, which is a non-increasing function of time tending to zero asymptotically.

$$R(t) = 1 - \sum_{k\in U} p_k(t) \tag{5.5}$$

[1] In literature there exist specific models for common mode failures, like the β model that splits the failure rate into two part, one common mode and one independent [39].

[2] An ordered statistics accounts for the timing when the events occur.

The Mean Time To Failure (MTTF) is calculated by the definition or with the final value theorem of the Laplace transform of $R(t)$:

$$MTTF = \lim_{t\to\infty} \int_0^t R(t)dt \overset{L}{\leftrightarrow} \lim_{s\to 0} R(s) \tag{5.6}$$

Any failure in the system may be recovered by means of repairs or scheduled inspections provided it has occurred in $X|U$.

5.2.1 Model without Repair

The unreliable set U can be reached through many different oriented acyclic paths[3] in X, starting from one initial state in $X|U$ and leading to one of the absorbing states in U. In the example of Figure 5.2(A) a generic path is shown, consisting of N + 1 states $(x_0, x_1, \ldots x_N)$, with $U \equiv x_N$, N transitions with constant failure rates λ_k, $(k = 0 \ldots N-1)$, and other N transitions with failure rate $\lambda_k^{'}$ that leave the path from each state x_k. The path is a realization of a stochastic process that can be modeled like an open Markov chain with state probability vector $p_k(t)$, $\{k = 0, 1, \ldots, N\}$ and initial distribution $\mathbf{p}(0) = [1, 0, \ldots, 0]$:

$$\frac{d}{dt}\mathbf{p}(t) = \mathbf{p}(t) \begin{pmatrix} -\lambda_0 - \lambda_0^{'} & \lambda_0 & 0 \cdots & 0 & 0 \\ 0 & -\lambda_1 - \lambda_1^{'} & \lambda_1 \cdots & 0 & 0 \\ \vdots & \vdots & \vdots \ldots & \vdots & \vdots \\ 0 & 0 & 0 \cdots & -\lambda_{N-1} & \lambda_{N-1} \\ 0 & 0 & 0 \cdots & 0 & 0 \end{pmatrix} \tag{5.7}$$

The probability of moving into the final state x_N and the MTTF associated to the path are respectively:

$$p_N(t) = 1 - \prod_{j=0}^{N-1} \frac{\lambda_j}{\lambda_j + \lambda_j^{'}} \int_0^t Hyp(\Lambda, \tau) d\tau \tag{5.8}$$

$$MTTF = \sum_{j=0}^{N-1} \frac{1}{\lambda_k + \lambda_k^{'}} \prod_{k=0}^{j} \frac{\lambda_k}{\lambda_k + \lambda_k^{'}} \tag{5.9}$$

[3] An oriented path is defined acyclic if the states are visited only once (no cycles).

where the function $Hyp(\Lambda, t) = Hyp(\lambda_0 + \lambda_0^{'}, \ldots, \lambda_{k-1} + \lambda_{k-1}^{'}, t)$ is the k stages hyperexponential distribution, resulting from the convolution of the probability distributions of the sojourn times of each state in the Markov chain, taken in the exact order they are visited. The system unreliability will be the sum of all contributions of the paths that lead to one state in U.

There exist several applications of this model in literature. As an example, the pure death process is obtained if $\lambda_j^{'} = 0, \forall j$, and reliability and MTTF are given by the following formulas:

$$R(t) = 1 - \sum_{i=0}^{N-1} \prod_{j=0, j \neq i}^{N-1} \frac{\lambda_j}{\lambda_j - \lambda_i} (1 - \exp(\lambda_i)) \tag{5.10}$$

$$MTTF = \sum_{i=0}^{N-1} \frac{1}{\lambda_i} \tag{5.11}$$

If all failure rates are identical, i.e. $\lambda_k = \lambda$, $\forall k = 0, \ldots N-1$, the probability to be in the state k is described by the Erlang distribution. The expression of reliability and MTTF are further simplified:

$$R(t) = \sum_{i=0}^{N-1} \frac{(\lambda t)^i}{i!} \exp(-\lambda t) = Erl(k, \lambda) \tag{5.12}$$

$$MTTF = N/\lambda \tag{5.13}$$

The model of Figure 5.2(B) is variant of the model given in 5.2(A), for which all output arcs lead to the failed state. This model can be used to the description of failures ($\lambda_k^{'}$) that lead straightway to the failed state (e.g. shock), and failures (λ_k) leading to progressive degeneration (e.g. wearing). The reliability at time t and the MTTF are:

$$R(t) = 1 - \sum_{k=0}^{N-1} \int_0^t \lambda_k^{'} p_k(t) dt - p_N(t) \tag{5.14}$$

$$MTTF = \overline{MTTF} + \frac{\lambda_0^{'}}{(\lambda_0 + \lambda_0^{'})^2} + \sum_{j=1}^{N-1} \frac{\lambda_j^{'}}{(\lambda_j + \lambda_j^{'})^2} \prod_{k=0}^{j-1} \frac{\lambda_k^{'}}{(\lambda_k + \lambda_k^{'})^2} \tag{5.15}$$

where $\overline{MTTF}$ is the MTTF calculated in (5.9).

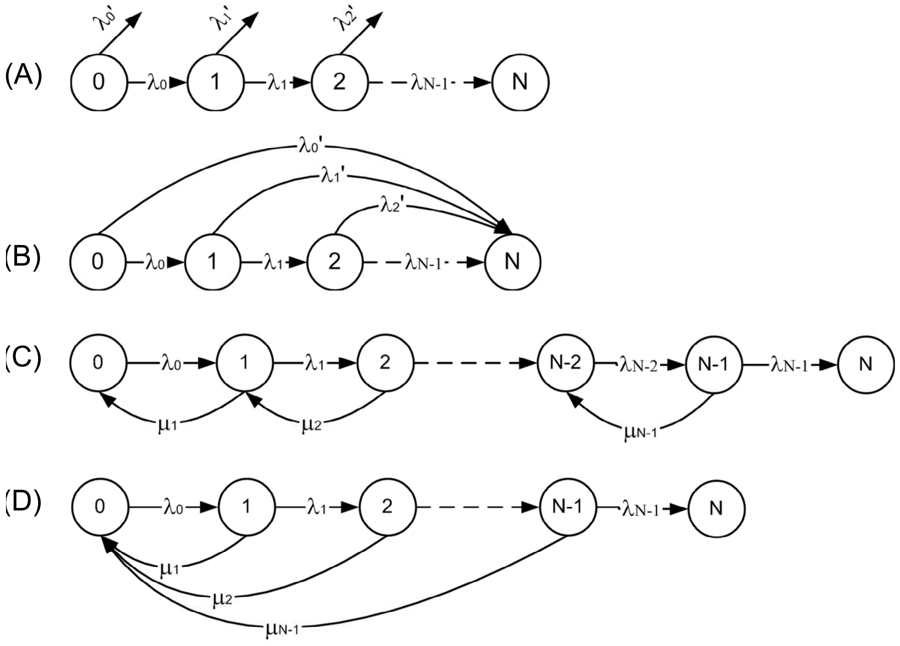


Figure 5.2: Models of failure processes.

5.2.2 Model with Repairs on Demand

Repair on demand recovers from failure in the subset $X|U$ just at the instant they are detected, with the effect of slowing down the failure process. As case study, a birth-death process with final absorbing state is considered, see Figure 5.2(C). The Kolgomorov equations that calculate the probability distribution are:

$$\frac{d}{dt}\mathbf{p}(t) = \mathbf{p}(t) \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \cdots & 0 & 0 \\ \mu_1 & -\lambda_1 - \mu_1 & \lambda_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\lambda_{N-1} - \mu_{N-1} & \lambda_{N-1} \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \tag{5.16}$$

for the initial state probability vector $\mathbf{p}(0) = [1, 0, \ldots, 0]$. The MTTF is calculated by applying the final value theorem of the Laplace transform to (5.16). The obtained formula is:

$$MTTF = \sum_{r=1}^{N-1} \sum_{k=0}^{N-r} \frac{1}{\lambda_k} \prod_{j=1}^{r-1} \frac{\mu_{k+j}}{\lambda_{k+j}} \tag{5.17}$$

A lower bound is $MTTF > \frac{1}{\lambda_{N-1}} \prod_{k=0}^{N-2} \frac{\mu_{k+1}}{\lambda_k}$, which still quantifies the benefits of the repair facilities. This model can be applied to describing a failure process where each state represent a recoverable degradation down to the complete system failure.

5.2.3 Model with Periodic Inspections

Periodic inspections permit to discover failures in the subset $X|U$ at scheduled instants, say every T. This model is shown in Figure 5.2(D), where the repair transitions are enabled at the time of inspection. The process is a Markov chain between two consecutive inspection intervals. At the time of inspection, the system is shutdown and all states, except for the final one, are recovered to the initial state. This has consequences on the system equivalent failure rate that becomes periodic with period T, with benefits on the failure process that slows down[4]. The reliability can be calculated with (5.10) within one inspection interval. After M consecutive inspections, during the mission M+1, the reliability and the MTTF are:

$$R(MT + t) = [1 - P_N(T)]^M [1 - P_N(t)] \tag{5.18}$$

$$MTTF = \frac{\int_0^T R(t)dt}{1 - R(T)} \tag{5.19}$$

where $0 \leq t < T$. A shorter inspection interval leads to a higher reliability, which is is generally true but does not take into account the inspection downtime. This is analysed more in detail in the next section.

5.3 Models for Availability

Availability models have the state space X of recurrent, non absorbing states. The unavailable set U is a subset of X. The metrics for availability are the instantaneous availability, the average availability and the steady state availability, also called Mean Time Between Failures (MTBF). The

[4] Partial inspections and imperfect repairs are not taken into account.

instantaneous availability $A(t)$ is the probability to be in the subset $X|U$ at time t. The average availability is the percent of time spent in $X|U$ during the interval T, $\bar{A}(T) = \frac{1}{T}\int_0^T A(t)dt$. The average availability tends asymptotically to the steady state availability[5], that is $A_{ss} = \lim_{t\to\infty} \bar{A}(t)$ [39].

5.3.1 Model with Repair on Demand

A model with repair on demand used for availability calculations is the model of Figure 5.2(C) provided that a repair transition is added from state x_N to x_{N-1}. A birth-death process of N +1 states is obtained, with U $\equiv$ N, constant failure rates λ_k and repair rates μ_k. The initial probability distribution is $\mathbf{p}(0) = [1, 0, \dots, 0]$. The Kolgomorov equations that calculate the probability distribution are:

$$\frac{d}{dt}\mathbf{p}(t) = \mathbf{p}(t)\begin{pmatrix} -\lambda_0 & \lambda_0 & 0 \cdots & 0 & 0 \\ \mu_1 & -\lambda_1 - \mu_1 & \lambda_1 \cdots & 0 & 0 \\ \vdots & \vdots & \vdots \dots & \vdots & \vdots \\ 0 & 0 & 0 \cdots & -\lambda_{N-1} - \mu_{N-1} & \lambda_{N-1} \\ 0 & 0 & 0 \cdots & \mu_N & -\mu_N \end{pmatrix} \tag{5.20}$$

The instantaneous availability is a non-increasing function of time t, tending asymptotically to the steady state value. A lower bound approximation for the steady state availability is:

$$A_{ss} = 1 - \lim_{t\to\infty} p_N(t) > 1 - \frac{1}{1 + \prod_{k=1}^{N} \frac{\mu_k}{\lambda_{k-1}}} \tag{5.21}$$

which quantifies the benefits of the repair facilities[6].

[5] For HCTMC this is obtained with the solution of Kolgomorov steady state equation $\mathbf{p}Q = 0$, where Q is the non singular transition matrix of the Markov chain [87].

[6] In the more general case, the model should account for a coverage factor C, representing the probability that the failure is detected.

5.3.2 Model with Periodic Inspections

An availability model with periodic inspection is shown in Figure 5.2(D) provided that a transition is added from the state x_N to x_0. The resulting instantaneous availability is periodic with period T. The steady state availability is the average availability for one inspection period T:

$$A_{ss} = \lim_{t \to \infty} \frac{\int_0^t [1 - P_N(t)]dt}{T} \tag{5.22}$$

The equation attests that a shorter inspection interval results in a higher availability. This conclusion does not take into consideration the total system downtime, which depends on the inspection interval T. If this is considered, then a trade-off can be found between availability and inspection policy. Assuming that D is the time necessary for the inspection, then the total system downtime at time t is:

$$T_D = \left\lfloor \frac{t}{T + D} \right\rfloor D \tag{5.23}$$

The shorter the inspection interval the bigger the downtime, which is the trade-off with the availability. A cost function of A_{ss} and T_D can be defined in order to find the optimal inspection interval T.

5.4 Models for Safety

Safety models have the state space X split into a set of safe states S and a set of unsafe states FU. The set S is further split into functioning state OP and failsafe state FS [63]. The safety is the probability to be in S. The definition applies to systems that operate either continuously (i.e. reliable) or on demand (i.e. available). The statistics for continuous safety are the MTTUF (Mean Time To Unsafe Failure), or MTTHE (Mean Time To Hazardous Event) [22], in analogy to the MTTF for reliability. In case the system operates on demand, the applied statistics are the safety at the steady state or the MTBUF (Mean Time Between Unsafe Failure), also defined like MTBUF = MTTUF + MTTR [86].

5.4.1 Model for non Recoverable Safety

The model for non recoverable safety is built on the simplified architecture for safety critical systems described in section 3.5 and consists of four states: the operating state OP1, the operating state without protection OP2, the failsafe state FS and the fail unsafe state FU, see Figure 5.3. The FU state is assumed to be absorbing, which makes safety to be non recoverable during mission time. For this reason, the model applies to those safety critical systems that operate continuously. The transition from OP1 to OP2, with rate λ_{P1}, represents the failure of the protection system. The transition from OP1 to FU, with rate λ_{S1}, represents the fraction of system failure uncovered by the protection system (e.g. the silent failure modes). The transition from OP2 to FU, with rate $\lambda_{S1} + \lambda_{S2}$, leads to the failure of the system when the protection system has failed. The fraction of the system failure rate λ_{S2} is now uncovered and adds up with its contribution. The transition from OP1 to FS, with rate $\lambda_{S2} + \lambda_{P2}$, is the fraction of system failure covered by the protection system plus the failures detected within the protection system.

After the system has failed safely at time T, this is inspected and recovered to the initial state. In case of perfect recovery, the system is 'as good as new' and $\lambda_k(T+) = \lambda_k(0)$ for all failure rates. The downtime for rearming the system does not concern safety and for this reason it is not considered. The resulting stochastic process is Markov regenerative with the failsafe events that govern the embedded Markov renewal sequence and represent the regeneration points of the process. The probability distribution is the solution of the Markov regenerative process with initial state probability distribution $p(0) = [1, 0, 0, 0]$, see section 4.4. The global kernel $K(t)$ and the local kernel $L(t)$ defined in equation (4.13) and (4.14) respectively are:

$$K(t) = \begin{pmatrix} -\lambda_{P2} - \lambda_{S2} & 0 & \lambda_{P2} + \lambda_{S2} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.24}$$

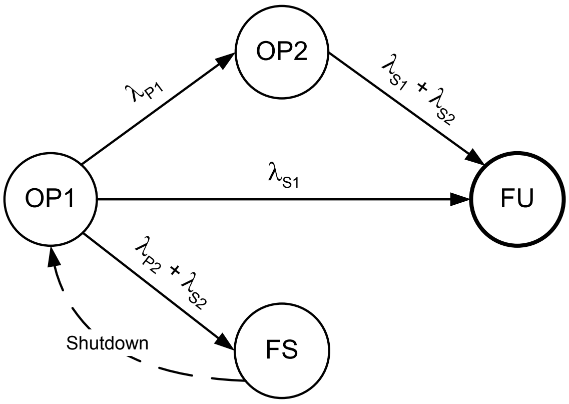


Figure 5.3: A safety model with failsafe shutdown.

$$L(t) = \begin{pmatrix} -\lambda_{P1} - \lambda_{S1} & \lambda_{P1} & 0 & \lambda_{S1} \\ 0 & -\lambda_{S1} - \lambda_{S2} & 0 & \lambda_{S1} + \lambda_{S2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.25}$$

The equation for safety is derived from equation (4.12):

$$S(t) = R(t) + S(t) \odot p_{FS}(t) \tag{5.26}$$

where $R(t) = p_{OP1}(t) + p_{OP2}(t)$, $S(0) = R(0) = 1$ and $\odot$ stands for the Laplace-Stieltjes convolution. The MTTUF is:

$$MTTUF = MTTF + \lim_{t\to\infty} \int_0^t S(t) \odot p_{FS}(t) \tag{5.27}$$

The safety $S(t)$ is the solution of a Volterra's second kind integral equation where the first term stands for the probability the system was continuously operating up to t (i.e. reliable) and the second term is the contribution of the failsafe shutdowns occurred in [0,t). The formulas (5.26) and (5.27) attest that safety is always bigger than reliability. The equation (5.26) can be transformed and solved in the Laplace domain:

$$S(s) = \frac{R(s)}{1 - sP_{FS}(s)} \tag{5.28}$$

As an alternative, the solution can be obtained by numerical integration.

The two following recursive formulas approximate the Laplace-Stieltjes convolution in (5.26) with the discrete summation [59]:

$$\hat{S}(t_n) = R(t_n) + \sum_{i=0}^{n} \hat{S}(t_n - t_i)[p_{FS}(t_i) - p_{FS}(t_{i-1})] \tag{5.29}$$

$$\check{S}(t_n) = \frac{R(t_n) + \sum_{i=1}^{n} \check{S}(t_n - t_{i+1})[p_{FS}(t_i) - p_{FS}(t_{i-1})]}{1 - p_{FS}(t_1)} \tag{5.30}$$

where $\hat{S}(t_0) = 1$, $\check{S}(t_0) = 1$, $R(t_0) = 1$, $p_{FS}(t_i) = 0$ for $t_i \leq 0$. Equation (5.30) is identical to equation (5.29) but with safety shifted in time of one discretization step. As this is a decreasing function with time, the (5.30) represents an upper bound of safety so that $\hat{S}(t_n) \leq S(t_n) \leq \check{S}(t_n)$. The arithmetic mean $(\hat{S}(t_n) + \check{S}(t_n))/2$ is an estimator of safety. The accuracy of the solution depends on the discretization step $t_i - t_{i-1}$. This aspect will be treated in Chapter 7 for a case study.

As an example, the transition rates of the Markov chain of Figure 5.3 are assumed to be constant: $\lambda_{S1} + \lambda_{S2} = \lambda_S$ and $\lambda_{S2} = C\lambda_S$, where the term C represents the fraction of failure rate covered by the protection system[7]. In this case the expression of the MTTUF becomes:

$$MTTUF = \frac{1}{\lambda_S} + \frac{C}{\lambda_{P1} + (1 - C)\lambda_S} \tag{5.31}$$

The maximum is obtained for $C = 1$ for which the MTTUF is the sum of the MTTF of the system and its protection system. In this case, a sufficient condition for the system to comply with the Safety Integrity Level x (x=1,2,3,4), see also section 3.5, is that the protection system is SILx. This is verified if $MTTUF \geq 1/\lambda_{SILx}$, where λ_{SILx} is the failure rate corresponding to SILx. If the protection system fails always safely (i.e. $\lambda_{P1} = 0$) then $MTTUF = MTTF/(1 - C)$ and the desired SIL depends only on the coverage C.

[7] Alternative definitions of coverage for safety can be found in [86] and in [47].

5.4.2 Model for Recoverable Safety

The model for recoverable safety includes a recovery transition from the FU state to OP1, which is not absorbing anymore, see Figure 5.4. The model applies to those systems operating on demand. The repair action is triggered by the detected fault, while the entire process is regenerated at the shutdown in the FS state. The failsafe shutdown still prevents from a system unsafe failure. The resulting stochastic process is Markov regenerative, with the failsafe events representing the regeneration points, and initial distribution $p(0) = [1, 0, 0, 0]$, see also section 4.4. $K(t)$ is given by equation (5.24) while $L(t)$ is:

$$L(t) = \begin{pmatrix} -\lambda_{P1} - \lambda_{S1} & \lambda_{P1} & 0 & \lambda_{S1} \\ 0 & -\lambda_{S1} - \lambda_{S2} & 0 & \lambda_{S1} + \lambda_{S2} \\ 0 & 0 & 0 & 0 \\ \mu & 0 & 0 & -\mu \end{pmatrix} \tag{5.32}$$

The instantaneous safety and the steady state safety are:

$$S(t) = A(t) + S(t) \odot p_{FS}(t) \tag{5.33}$$

$$S_{ss} = A_{ss} + \lim_{t \to \infty} \frac{\int_0^t S(t) \odot p_{FS}(t)}{t} \tag{5.34}$$

where $A(t) = p_{OP1}(t) + p_{OP2}(t)$ and $S(0) = A(0) = 1$. The formulas attest that safety is bigger than availability. The equation (5.33) can be transformed and solved in the Laplace domain:

$$S(s) = \frac{A(s)}{1 - sP_{FS}(s)} \tag{5.35}$$

As an alternative, the solution can be obtained numerically like in (5.29) and (5.30) provided that $R(t)$ is replaced by $A(t)$.

Assuming that the failure rates are constant with $\lambda_{S1} + \lambda_{S2} = \lambda_S$ and $\lambda_{S2} = C\lambda_S$, where C is the fraction of failure rate covered by the protection system, the expression for safety at the steady state becomes:

$$S_{ss} = 1 - \frac{(1-C)\lambda_S^2 + \lambda S \lambda_{P1}}{(1-C)\lambda_S^2 + \lambda_{P1}\mu + \lambda_S(\mu + \lambda_{P1})} \tag{5.36}$$

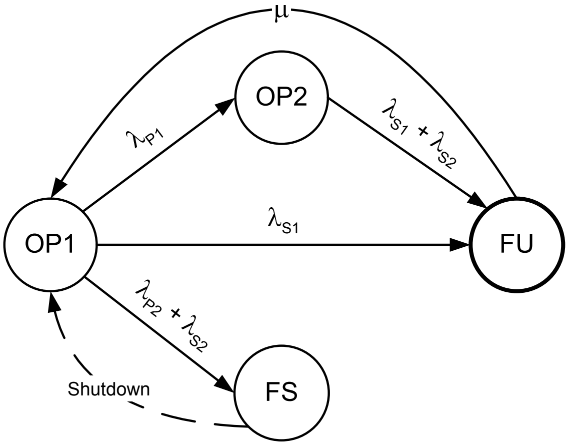


Figure 5.4: A safety model with failsafe shutdown and repair.

For C = 0, the protection system does not play any role. The resulting safety at the steady state is $S_{ss} = 1 - \frac{\lambda_S}{\lambda_S+\mu}$, which ameliorates by increasing the repair rate. If C = 1 then $S_{ss} = 1 - (\frac{\mu}{\lambda_S} + \frac{\mu}{\lambda_{P1}} + 1)^{-1}$. In case the protection system never fails silent (i.e. $\lambda_{P1} = 0$), then $S_{ss} = 1 - (1 + \frac{\mu}{(1-C)\lambda_S})^{-1}$ and the coverage and the repair rate are the parameters to reach the required SIL[8].

A variant of the illustrated model includes periodic inspections at every fixed interval T instead of repair on demand, in addition to the failsafe shutdown, see Figure 5.5. The inspection is modeled with a transition from states OP2 and FU to OP1, which takes place at the instant T. If the system moves to FU, it must wait for the next inspection as the failsafe shutdown is inhibited in FU. After inspection the system is assumed to be recovered 'as good as new'. The resulting stochastic process is Markov regenerative, with the failsafe shutdown and the periodic inspections acting as regeneration points. The solution is calculated for one interval T, during which the system may undergo a certain number of safe shutdowns. The instantaneous safety is periodic with period T, that is $S(t+T) = S(t)$, with $S(t)$ obtained from equation (5.26). Also the equivalent failure rate to the unsafe state is periodic with period T. Safety at steady state is the average of $S(t)$ over the

[8] More specialized case studies can be found in [14, 36].

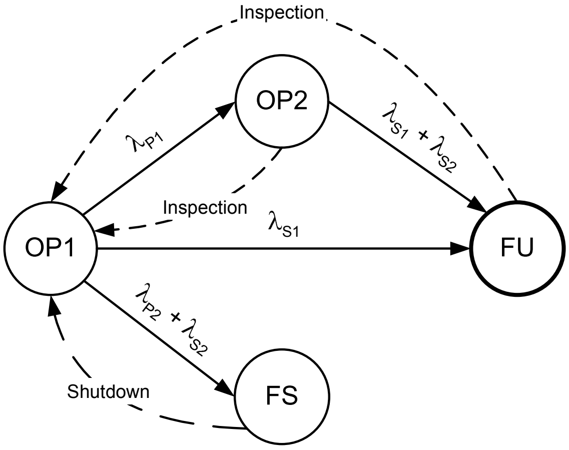


Figure 5.5: A safety model with failsafe shutdown and periodic inspection.

time interval T between two consecutive inspections:

$$S_{ss} = \frac{1}{T}\frac{C(1-e^{-[(1-C)\lambda_S+\lambda_{P1}]T})\lambda_S}{[(1-C)\lambda_S+\lambda_{P1}](C\lambda_S-\lambda_{P1})} + \frac{1}{T}\frac{(1-e^{-\lambda_S T})\lambda_{P1}}{\lambda_S(-C\lambda_S+\lambda_{P1})} \tag{5.37}$$

If coverage C = 0 then $S_{ss} = \frac{1}{T}(1-e^{-\lambda_S T})\lambda_S^{-1}$. If C = 1, then $S_{ss} = \frac{1}{T}\frac{C(1-e^{-\lambda_{P1}T})\lambda_S}{\lambda_{P1}(\lambda_S-\lambda_{P1})} + \frac{1}{T}\frac{(1-e^{-\lambda_S T})\lambda_{P1}}{\lambda_S(-\lambda_S+\lambda_{P1})}$.

In case FU is absorbing, the safety at time t, as calculated by (5.26), and the MTTFU become:

$$S(t+MT) = S(T)^M S(t) \tag{5.38}$$

$$MTTUF = \frac{\int_0^T S(t)dt}{1-S(T)} \tag{5.39}$$

The downtime due to inspection interval, presented in section 5.3.2 for availability, is a trade-off with safety too.

5.4.3 Other Models for Safety

Other models for the analysis of safety critical systems exist. One of these represents the status of the protection system embedded in the transition rates [47], see Figure 5.6 (right). This model is equivalent to the model of Figure 5.3. Three states are obtained OP1, FS, FU instead of four, with transition rates $\lambda_{OP1-FS} = (\lambda_{P2}+\lambda_{S2})R_P(t)$ and $\lambda_{OP1-FU} = \lambda_{S1} + \lambda_{S2}[1-R_P(t)]$,

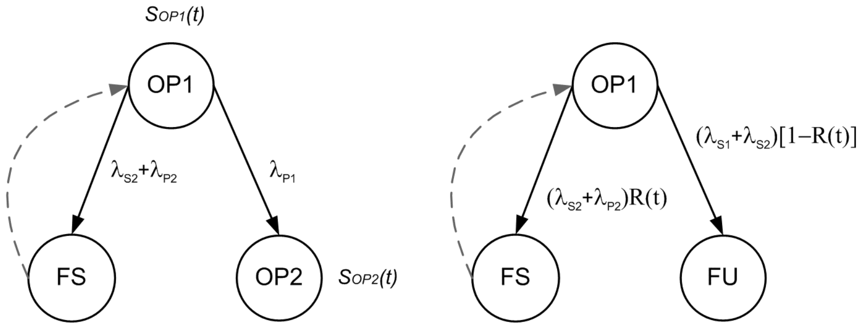

$S_{OP1}(t)$
OP1
$\lambda_{S2}+\lambda_{P2}$
λ_{P1}
FS
OP2
$S_{OP2}(t)$
OP1
$(\lambda_{S1}+\lambda_{S2})[1-R(t)]$
$(\lambda_{S2}+\lambda_{P2})R(t)$
FS
FU

Figure 5.6: Alternative models for safety.

where $R_P(t)$ is the reliability of the protection system that depends on λ_{P1}. Safety becomes:

$$S(t) = p_{OP1}(t) + S(t) \odot p_{FS}(t) \tag{5.40}$$

which is solved for the initial conditions $S(0) = P_{OP1}(t) = 1$.

Another model represents the status of the protection equipment instead of the system, see Figure 5.6 (left). Three states exist, OP1, OP2, and FS with transition rates $\lambda_{OP1-OP2} = \lambda_{P1}$ and $\lambda_{OP1-FS} = \lambda_{P2} + \lambda_{S2}$. The system safety is a cumulative reward function of the three states defined on the Markov regenerative reward process, see section 4.5:

$$S(t) = S_{OP1}(t)p_{OP1}(t) + S_{OP2}(t)p_{OP2}(t) + S(t) \odot p_{FS}(t) \tag{5.41}$$

which is solved for the initial conditions $p_{OP1}(0) = 1$ and $p_{OP2}(0) = p_{FS}(0) = 0$. The functions S_{OP1}, and S_{OP2} are the reward functions in the states OP1 and OP2 respectively. The model is equivalent to the models in Figures 5.6 (right) and 5.3.

It is possible that more levels of protection exist for a safety critical system. For example, consider K protection systems, each one covering a fraction of the system failure rate. The exact model will account for $2^K + 2$ states. A fair approximation is to disregard the second order failures (2 or more protection systems failed) provided that these are independent, see Figure 5.7. The approximated model will consist of $K + 3$ states. The

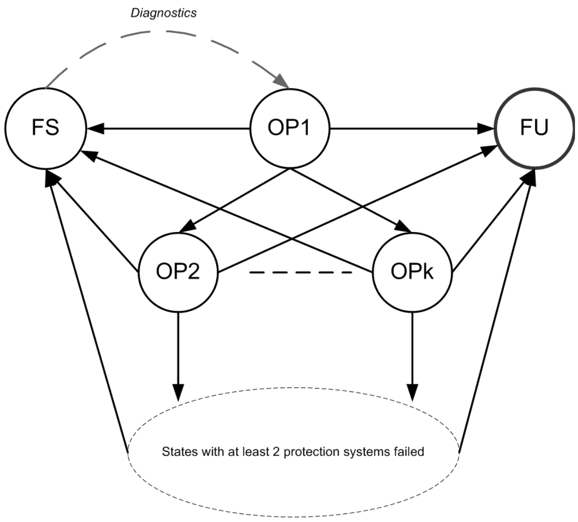


Figure 5.7: A safety model for more protection systems.

calculated safety will be an upper bound, suited for checking the compliance with SIL level with some tolerance margin.

All presented models have the common feature of issuing an operation shutdown in case of a failure has been detected. Actually, the decision taken at the occurrence of a detected failure might not necessarily be a shutdown. For instance, it could depend on the type of failure, the overall status of the system and the phase criticality. To include fault management and supervision in the models would lead to a Markov decision process and is not treated here.

5.4.4 The Safety Trade-off

The mission shutdown stops the operation at a possible detriment of the system reliability or availability. This represents a trade-off for the reachable safety especially if caused by false alarms within the protection system. The number of operation aborts due to system shutdowns is calculated in the same way for all presented models. At time t, the system will have undergone N operation shutdowns, part of these due to detected system failures (N_{S2})

and part due to failures in the protection system (N_{P2}). The number of operation shutdowns $N(t)$ is the result of a renewal counting process with renewal density $\lambda_{P2} + \lambda_{S2}$, see equation (4.7). The fundamental renewal equation provides the average $M(t) = E[N(t)]$:

$$M(t) = p_{FS}(t) + M(t) \odot p_{FS}(t) \tag{5.42}$$

for the initial conditions $M(0) = p_{FS} = 0$, where $\odot$ stands for the Laplace-Stieltjes convolution. The solution is obtained by Laplace transform or by numerical integration like in (5.29) and (5.30). The upper and lower bound expressions for the number of missions at time t are:

$$\hat{M}(t_n) = p_4(t_n) + \sum_{i=0}^{n} \hat{M}(t_n - t_i)[p_4(t_i) - p_4(t_{i-1})] \tag{5.43}$$

$$\check{M}(t_n) = \frac{p_4(t_n) + \sum_{i=1}^{n} \check{M}(t_n - t_{i+1})[p_4(t_i) - p_4(t_{i-1})]}{1 - p_4(t_1)} \tag{5.44}$$

where $\hat{M}(0) = 0$, $\check{M}(0) = 0$, $p_{FS}(t_i) = 0$ for $t_i \leq 0$ and $\hat{M}(t_n) \leq M(t_n) \leq \check{M}(t_n)$ for the same reasons given for the calculation of safety. The arithmetic mean $(\hat{M}(t_n) + \check{M}(t_n))/2$ is an estimator of $M(t)$, which depends on the discretization step $t_i - t_{i-1}$.

If the failure rates are assumed to be constant, then the distribution of $N(t)$ is a Poisson process:

$$P(N(t) = n) = \frac{(\lambda_{P2} + \lambda_{S2})^n}{n!} e^{-(\lambda_{P2}+\lambda_{S2})t} \tag{5.45}$$

for which the average number of operation aborts is $E[N, T] = (\lambda_{P2} + \lambda_{S2})t$ and the variance is $\lambda_{P2} + \lambda_{S2}$. The conditional distribution of the operation shutdowns due to false alarms N_{P2} is:

$$P(N_{P2}(t) = n_{P2} | N(t) = n) = \binom{n}{n_P} (\frac{\lambda_{P2}}{\lambda_{P2} + \lambda_{S2}})^{n_P} (\frac{\lambda_{S2}}{\lambda_{P2} + \lambda_{S2}})^{n-n_P} \tag{5.46}$$

for which the average number of false alarms is $E[N, T] = \lambda_{P2} t$. The above formula does not hold anymore for non constant failure rates. In this case other quantities need to be taken into consideration, for example the average of the failure rates over a certain time interval.

The trade-off existing between safety and the number of operation aborts, due to false alarms, can be formulated like an optimization problem. As an alternative, a qualitative approach is followed. If operation aborts due to false alarms exceed an acceptable limit (e.g., 30%), a reduction can be reached through voting mechanisms and signal/data processing that filter out the undesired shutdown. As drawback this solution adds complexity to the system, which may introduce undesired factors of uncertainty for the safety assessment, see also 3.4.5. On the contrary, if the largest contribution to the operation aborts is due to the system failure rate λ_{S2}, then system reliability becomes the main concern and a design revision is to be taken in consideration.

Chapter 6

Failure Modes Analysis of the LHC Beam Dumping System

This chapter describes in detail the failure mechanisms of the LHC Beam Dumping Systems LBDS ad deduced from the architecture of the system presented in Chapter 2.

6.1 The Modeling Framework

The dependability study of the LBDS concerns those systems that play an active role in the removal and dilution of the beams from the ring. This is the core-architecture of the LBDS. They are the MKD, the MSD, the MKB, the triggering system and the Beam Energy Measurement System (BEMS), plus the Beam Energy Tracking System (BETS) and the Re-Triggering System (RTS), see Figure 6.2. The quadrupole Q4 is not included in the core architecture because its failure is detected independently from the LBDS, for instance as beam loss by the beam loss monitors or by the quench protection system. The TDE, the TCDS, the TDCQ and the TCS as well as the beam instrumentation are also outside the core architecture. In case of malfunctioning of any of these systems, the LBDS becomes unavailable with possible consequences on safety.

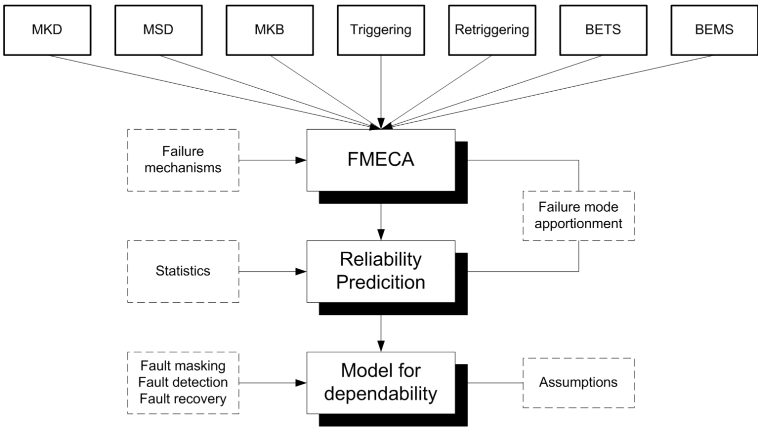


Figure 6.1: Dependability modeling and analysis framework.

The structure of the study is shown in Figure 1. The adopted approach combines qualitative (FMECA [65, 67, 34]) and quantitative techniques (reliability prediction [66]), applied at lower level, and the state based representation of the system failure process at a higher level. The modeling hierarchy permits to mask details, while working on the higher level, and gives the possibility to analyse the system for a large set of parameters, different operational scenarios and the design features like fault masking (i.e. redundancy), fault detection (on-line), post mortem diagnostics and fault recovery (off-line). This chapter contains the FMECA of the core LBDS and its sub-systems, while the reliability prediction and the dependability analysis of the model of the system failure process will be treated in the next Chapter.

6.2 The LBDS Failure Modes

Most malfunctioning in the LBDS does not lead to beam losses or leads to beam losses that can be tolerated as their consequences are mitigated by the passive protection devices. Some ‘beyond design’ failures exist and may lead to the loss of the entire beam with catastrophic consequences [89]. This is the case of failures during the beam extraction of the MKD or the MSD magnet

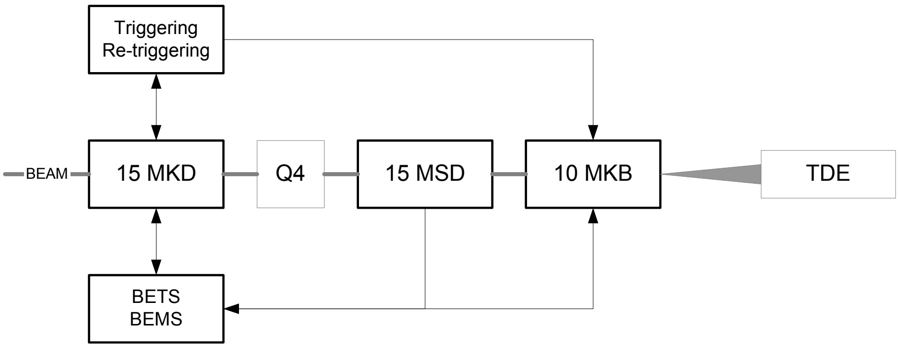


Figure 6.2: LBDS functional description.

assemblies or in the control electronics like the beam energy measurement and tracking and the triggering system. Failures in the beam dilution are less critical though the complete unavailability of the MKB system might destroy the dump block, requiring long downtime for repair or replacement.

The system failure modes can be classified in three categories: fail silent modes, failsafe system failures modes and failsafe surveillance failures modes, which are false alarms.

- F1: fail silent and undetectable.
- F2: failsafe due to detectable faults in the system.
- F3: failsafe due to false alarms in the surveillance devices.

The F1 failure mode affects safety. Fault tolerance may mask to a certain extent these type of failure modes that accumulate undetected before the system fails silent, and can only be discovered when the LBDS is in the firing mode, see Chapter 2. The F2 failure mode affects both safety and availability. Surveillance cover these failure modes generating an operation abort if they occur. In case the surveillance has failed, these failure modes turn to be silent. The F3 failure modes affects availability and not safety, as they do not correspond to any hazard in the machine. Safety and availability are a trade-off for the LBDS as demonstrated in 5.4.4: for a higher level of safety the availability is reduced. The maximum safety level is reached if the number of F1 failure modes is reduced by system design and all remaining F2 failure modes are detected (i.e. the coverage is unitary) while the minimum safety is reached if the F2 modes are missed and add to F1 (i.e. the coverage tends to zero).

A complete list of failure modes of the core LBDS architecture of Figure 6.2, taken at system level, is summarized in Table 6.1. They are classified with respect to the three modes F1, F2 and F3, the detection (if any) and the consequences on the dependability attributes (i.e. safety and availability) according to a FMECA-like approach presented in section 3.2. For the small part of failsafe modes that are self-announcing the dump request is automatically generated. Surveillance devices like the BETS and the RTS are listed separately in Table 6.1, at the same level of the other systems, while other internal surveillance is apportioned to each system. This choice is made on the basis of the importance of the monitoring of energy tracking failures and erratic triggers. The naming of the failure modes includes the type of failure, the system where it occurs and the position in the Table. For example, $F2_{MKD1}$ refers to the first failure mode of type F2 that is listed for the MKD system. This failure mode has partial internal compensation by redundancy, it is covered by the beam energy tracking system and may have consequences on safety and availability.

The following assumptions have been made throughout the FMECA:

- Assumption 1: All failure modes are assumed to be confined to the component where they develop, ignoring possible fault propagations.
- Assumption 2: All failures are assumed concurring and statistically independent events, see section 5.1.
- Assumption 3: All failsafe modes, once triggered, are assumed to lead to a safe state.

The first assumption makes it possible to build a system hierarchy where each component may be separately analysed at the required level of detail. As an example, a trigger failure is caused by the triggering system in case this affect all MKD and MKB generators or an asynchronous trigger has been generated. A trigger failure can develop within the MKD system if the switching mechanism inside a generator did not function or one erratic trigger has been generated. Analogously, the energy tracking failures is caused by the BEMS, where the beam energy is calculated and distributed, and affects all generators, or it may develop within each generator. The second assumption take some benefits in the composition of the logic expression of failure while the third neglects the possibility the dump request is generated when the system had already failed, which is unlikely and already accounted for as a system failure. About the mutually exclusiveness of failures, it is possible that some failure may accumulate while others cannot, and their occurrence may inhibit the occurrence of other failure modes, see also section 5.1. In the present description, this level of detail is ignored, as the gain in accuracy would not justify the added mathematical complexity.

The failure modes in Table 6.1 will be separately deduced for each system in the next sections.

6.3 The MKD System Failure Modes

The detailed list of failure modes of the MKD components is shown in Table 6.2. These failure modes combine together in logic expressions that describe the failure modes at system level, given in Table 6.1, applying the logic operators in Table 6.3. The five failure modes of the MKD system are:

$F1_{MKD}$: **two or more MKD systems have failed silent**. The identified causes for this failure mode are: the power triggers that are not responding to the input trigger (PT_1, see Table 6.2), the primary or the compensation switches that are not closing (SP_1, SC_1), the primary capacitors or the overshoot capacitors charged but not connected to the circuit (CP_2, $COS1_2$, $COS2_2$) and the magnet failure including the transmission cables and the connections (M). These failure modes are arranged in the following

Failure modes		**Compensation**	**Detection**	**Attribute**
MKD system				
$F1_{MKD}$	Less that 14 MKD available	Redundancy	No	Safety
$F2_{MKD1}$	Energy tracking failure	Partial Red.	BETS	Safety/Av.
$F2_{MKD2}$	Erratic trigger	No	RTS	Safety/Av.
$F2_{MKD3}$	Power supplies failures	No	IS	Availability
$F3_{MKD}$	IS false alarm	No	Self-an.	Availability
MSD system				
$F2_{MSD1}$	Energy tracking failure	No	BETS	Safety
$F2_{MSD2}$	Fast current changes	No	IS	Safety/Av.
$F3_{MSD}$	IS false alarm	No	Self-an.	Availability
MKB system				
$F1_{MKB}$	No MKBH or no MKBV available	Redundancy	No	Safety
$F2_{MKB1}$	Energy tracking failure	Partial Red.	BETS	Safety/Av.
$F2_{MKB2}$	Power supplies failures	No	IS	Availability
$F3_{MKB}$	IS false alarm	No	Self-an.	Availability
Triggering system (trigger generation and distribution)				
$F1_{TRG}$	No trigger	Redundancy	No	Safety
$F2_{TRG1}$	Spurious triggers	No	Self-an.	Availability
$F2_{TRG2}$	Synchronization error	Redundancy	IS	Availability
$F3_{TRG}$	IS false alarm	No	Self-an.	Availability
Beam Energy Meter System (BEMS)				
$F2_{BEMS}$	Energy tracking failure	Redundancy	IS, BETS	Safety/Av.
$F3_{BEMS}$	IS false alarm	No	Self-an.	Availability
Beam Energy Tracking System (BETS)				
$F1_{BETS}$	Unable to trigger a dump request	No	No	Safety (coverage)
$F3_{BETS}$	False alarm	No	Self-an.	Availability
Re-Triggering System (RTS)				
$F1_{RTS}$	Unable to re-trigger	Redundancy	No	Safety (coverage)

Table 6.1: Failure modes of the LHC Beam Dumping System.

Failure modes		**Type**	**Compensation**	**Detection**
Power supplies				
PSP_1	Primary PS under-voltage	Failsafe	No	BEI, BETS
PSP_2	Primary PS over-voltage	Failsafe	No	BEI, BETS
$PSOS_1$	Overshoot1 PS failure	Failsafe	No	BEI, BETS
$PSOS_2$	Overshoot2 PS failure	Failsafe	No	BEI, BETS
Power triggers				
PT_1	Power trigger not responding	Fail silent	Redundancy	Diagnostics
PT_2	Erratic trigger	Failsafe	No	Re-triggering
PT_3	Power supply failure	Failsafe	No	IS
Capacitors				
CP_1	Primary capacitor slow leakage	Failsafe	No	BEI, BETS
CP_2	Primary capacitor open failure	Fail silent	Redundancy	Diagnostics
$COS1_1$	Overshoot1 capacitor slow leakage	Failsafe	No	BEI, BETS
$COS1_2$	Overshoot1 capacitor open failure	Fail silent	Redundancy	Diagnostics
$COS2_1$	Overshoot2 capacitor slow leakage	Failsafe	No	BEI, BETS
$COS2_2$	Overshoot2 capacitor open failure	Fail silent	Redundancy	Diagnostics
Switches				
SP_1	Primary s. fails to open	Fail silent	Redundancy	Diagnostics
SP_2	Primary s. closes erratically	Failsafe	No	Re-triggering
SC_1	Compensation s. fails to open	Fail silent	Redundancy	Diagnostics
SC_2	Compensation switch closes erratically	Failsafe	No	Re-triggering
Magnet				
M	Magnet failure	Fail silent	No	Diagnostics
Beam energy data acquisition ad interlocking				
VD	Voltage divider failure	Failsafe	No	BEI, BETS
BEA	Data acquisition error	Failsafe	No	BEI, BETS
BEI_1	BEI unavailable	Fail silent	No	Diagnostics
BEI_2	False alarm	Failsafe	No	Self-an.

Table 6.2: MKD system failure modes.

Operator	**Symbol**
AND	$\wedge$
OR	$\vee$
NOT	$\neg$
XOR	$\dot{\vee}$
k out of n	$\binom{n}{k}$
AND T-priority	$\wedge_T$

Table 6.3: Logic operators.

logic expression for $F1_{MKD}$:

$$F1_{MKD} = \binom{15}{2}[MKD_{silent}] \tag{6.1}$$

$$\begin{gathered} MKD_{silent} = (PT_{1A} \wedge PT_{1B}) \vee (SP_{1A} \wedge SP_{1B}) \vee (SC_{1A} \wedge SC_{1B}) \vee (CP_{2A} \wedge CP_{2B}) \vee \\ (COS1_{2A} \wedge COS1_{2B}) \vee (COS2_{2A} \wedge COS2_{2B}) \vee M \end{gathered}$$

A and B are the generator branches.

$F2_{MKD1}$**: the system has failed due to an energy tracking failure**. This failure mode is surveyed by the Beam Energy Interlocking (BEI) cards, one per MKD generator, 15 in total. The BEI generates locally the dump request that sends to the BETS[1]. If all 15 BEI and the BETS are functioning, the failure mode is fully covered and detectable otherwise pert of the system is exposed to an increased hazard of failing undetected. The logic expression for the undetected failure mode is the following[2]:

$$\widetilde{F2}_{MKD1} = \dot{\vee}_{k=1\ldots15}[X_{BEI}(k) \wedge MKD_{uncovered}(k)] \tag{6.2}$$

$$MKD_{uncovered}(k) = (\tbinom{k}{1}[MKD_{energy}] \wedge \tbinom{14}{1!}[MKD_{silent}]) \vee \tbinom{k}{2}[MKD_{energy}] \vee \tbinom{k}{1}[PSP_2]$$

$$\begin{gathered} MKD_{energy} = \\ (PSP_1 \vee PSOS_1 \vee PSOS_2 \vee CP_{1A} \vee CP_{1B} \vee COS1_{1A} \vee COS1_{1B} \vee COS2_{1A} \vee COS2_{1B}) \end{gathered}$$

The variable X_{BEI}(k) means that k BEI cards have failed silent (BEI_1) and 15-k are functioning, that is $\neg(BEI_1 \vee BEI_2)$. All energy-tracking failures are masked by 14 out of 15 redundancy with the exception of the over-voltage of the power generator (PSP_2). The operator $\binom{k}{1}[MKD_{energy}]$ means that only one of the k uncovered magnets had an energy tracking failure, which is not a failure for the system but, in combination with exactly one silent failure in the remaining 14 magnets $\binom{14}{1!}[MKD_{silent}]$, is another way of failing[3]. If

[1] The analysis of failure modes of the BEI and the BETS is done in section 6.9.

[2] Hereafter, the superscript~is used for the undetected failsafe mode while the superscript^is used for the detected ones.

[3] This notation means that only one failure out of 14 is taken. The occurrence of two of them is a system failure and it is already accounted in $F1_{MKD}$.

the BETS has failed silent, all BEI are blind and the failure mode expression becomes:

$$\widetilde{F2}_{MKD1noBETS} = MKD_{uncovered}(15) \tag{6.3}$$

The failure mode is detected in case that it has occurred within the set of the functioning BEI cards. The logic expression becomes:

$$\widehat{F2}_{MKD1} = \dot{\vee}_{k=1\ldots14}[X_{BEI}(k) \wedge MKD_{covered}(k)] \tag{6.4}$$

$$MKD_{covered}(k) = \tbinom{(15-k)}{1}[MKD_{energy} \vee PSP_2] \vee \tbinom{k}{2}[MKD_{energy}]$$
$$\wedge \tbinom{14}{1!}[MKD_{silent}]$$

Again, this failure mode is not detectable if the BETS has failed silent.

$F2_{MKD2}$: **The system has failed due to an erratic trigger**. This failure mode is surveyed by the Re-Triggering System (RTS). It occurs if at least one erratic has been generated in either one power trigger (PT_2) or one switch (SP_2, SC_2) in at least one of the 30 generator branches A and B per LBDS system. The risk of leaving an erratic trigger uncovered at the source is largely dominated by the probability that both the re-triggering lines have failed, as it will be demonstrated in this Chapter, section 6.7. The logic expression for $F2_{MKD2}$ is:

$$F2_{MKD2} = \binom{15}{1}[PT_{2A} \vee PT_{2B} \vee SP_{2A} \vee SP_{2B} \vee SC_{2A} \vee SC_{2B}] \tag{6.5}$$

$F2_{MKD3}$: **The system has failed due to a failure in one power trigger power supply**. This failure mode is generated by the detected failure of at least one power supply in the 30 power triggers of the MKD system. The logic expression is:

$$F2_{MKD3} = \binom{15}{1}[PT_{3A} \vee PT_{3B}] \tag{6.6}$$

$F3_{MKD}$: **The surveillance has generated a false alarm**. The system fails safely for a false alarm if the BEI has internally generated a false alarm or the beam energy acquisition, including the BEA[4] and the voltage divider,

[4] The failure modes analysis of the BEA is done in section 6.8.

Failure modes		**Type**	**Compensation**	**Detection**
Driver primary				
DP_1	Fail to trigger	Fail silent	Redundancy	Diagnostics
DP_2	The driver fires erratically	Failsafe	No	Re-triggering
Switch primary				
$PTSP_1$	The switch does not close	Fail silent	Redundancy	Diagnostics
$PTSP_2$	The switch closes erratically	Failsafe	No	Re-triggering
Driver compensation				
DC_1	Fail to trigger	Fail silent	Redundancy	Diagnostics
DC_2	The driver fires erratically	Failsafe	No	Re-triggering
Switch compensation				
$PTSC_1$	The switch does not close	Fail silent	Redundancy	Diagnostics
$PTSC_2$	The switch closes erratically	Failsafe	No	Re-triggering
Redundant trigger path				
RP_1	Failed open	Fail silent	Redundancy	Diagnostics
RP_2	Erratic trigger	Failsafe	No	Re-triggering
Power supplies				
PTM	Breakdown	Failsafe	No	IS
HV	Breakdown	Failsafe	No	IS
PTC	Breakdown	Failsafe	No	IS
Power Trigger Controller				
PTC_1	Surveillance unavailable	F. Silent	No	Diagnostics
PTC_2	False alarm	Failsafe	No	Diagnostics
PTC_3	Erroneous ref. voltage	Fail silent	No	Diagnostics

Table 6.4: Power trigger failure modes.

has failed. This latter failure assumes the availability of the BEI. The logic expression is:

$$F3_{MKD} = \binom{15}{1}[BEI_2 \vee [(BEA \vee VD) \wedge \neg(BEI_1 \vee BEI_2)]] \tag{6.7}$$

6.3.1 The Power Triggers Failure Modes

The failure modes of the power trigger are shown in Table 6.4. They combine together to obtain the failure modes PT_1, PT_2 and PT_3, which enter the expressions of some MKD and MKB failure modes.

PT_1: **The system is not able to trigger**. This failure mode occurs if the primary or the compensation circuits have failed in both modules $(DP_1, RP_1, PTSP_1, DC_1, PTSC_1)$ or these are powered at the wrong voltage setting (PTC_3). In addition to this, there is the contribution of the undetected failure of the power-supplies (PTM-PS, HV-PS), which may occur if the Power Trigger Controller PTC has failed silent (PTC_1) before the failure of the power supply (i.e. time priority). The logic expression is:

$$PT_1 = PT_{1A} \vee PT_{1B} \vee PTC_3 \vee [(PTM \vee HV) \wedge_T (PTC_1)] \qquad (6.8)$$

$$PT_{1A/B} = [(DP_{1A/B} \wedge RP_{1A/B}) \vee PTSP_{1A/B} \wedge (DC_{1A/B} \vee PTSC_{1A/B})]$$

PT_2: **The system has generated an erratic trigger**. An erratic trigger may occur either in the primary or in the compensation circuit. The logic expression is:

$$PT_2 = PT_{2A} \vee PT_{2B} \qquad (6.9)$$

$$PT_2 = DP_{2A/B} \vee RP_{2A/B} \vee PTSP_{2A/B} \vee DC_{2A/B} \vee PTSC_{2A/B}$$

PT_3: **The system has a power supply failure**. A safe dump request is generated when the PTC has detected a failure in the power supplies of the PTC, the PTM and the HV capacitor.

$$PT_3 = [(PTC \vee PTM \vee HV) \wedge (X_{PTC})] \vee PTC_2 \qquad (6.10)$$

This failure is detected if $X_{PTC} = \neg(PTC_1 \vee PTC_2 \vee PTC_3)$.

6.4 The MSD System Failure Modes

The failure modes for the extraction septa MSD are listed in Table 6.5. They combine together in the system failure modes listed in Table 6.1. Due to continuous surveillance of the status of the system, only failsafe modes of type F2 and F3 exist.

$F2_{MSD1}$: **The system has failed due to an energy tracking failure**. The failure mode depends on the status of the local BEI card that is expressed with the variable X_{BEI}. The logic expression is:

$$F2_{MSD1} = PC_1 \wedge X_{BEI} \qquad (6.11)$$

Failure modes		**Type**	**Compensation**	**Detection**
Power converter				
PC_1	Slow failure	Failsafe	No	BEI, BETS
PC_2	Fast failure/power cut-off	Failsafe	No	FMCCM
Magnet				
M_1	Field out of tolerance	Failsafe	No	FMCCM
M_2	Shorts developed slowly	Fail silent	No	Diagnostics
PLC and thermo-switches				
PLC	PLC unavailable	Fail silent	BEI, FMCCM	Diagnostics
TS_1	Failed stuck-at	Fail silent	Sec. effects	Diagnostics
TS_2	Erratic trigger	Failsafe	No	PLC
FMCCM				
FC_1	Unavailable	Fail silent	No	Diagnostics
FC_2	The switch closes erratically	Failsafe	No	Re-triggering
VD	Generic failure	Failsafe	No	FMCCM
Beam Energy Data Acquisition and Interlocking				
$DCCT$	Generic failure	Failsafe	No	BEI, BETS
BEA	Error in data acquisition	Failsafe	No	BEI, BETS
BEI_1	Unavailable	Fail silent	No	Diagnostics
BEI_2	False alarm	Failsafe	No	Self announcing

Table 6.5: MSD system failure modes.

The failure is undetected if $X_{BEI} = BEI_1$ and the event occurred before PC_1 (i.e. time priority). This is detected if $X_{BEI} = \neg(BEI_1 \vee BEI_2)$. In the case the BETS has failed silent, whatever the state of the BEI card, the failure mode PC_1 remains undetected.

$F2_{MSD2}$: **The system has failed due to a fast load change**. The failure mode depends on the status of the FMCCM:

$$F2_{MSD2} = (PC_2 \vee \binom{15}{1}[M_1]) \wedge X_{FMCCM} \tag{6.12}$$

The failure is undetected if $X_{FMCCM} = FC_1$ and the event occurred before the failure (i.e. time priority). This is detected if $X_{FMCCM} = \neg(FC_1 \vee FC_2)$.

$F3_{MSD}$: **The surveillance has produced a false alarm**. The logic expression for this failure mode accounts for failures in the BEA-BEI, in the FMCCM, in the voltage divider of the FMCCM, in the DCCT magnet current measurement and in any failure of the thermo-switches that may lead

Failure modes		**Type**	**Compensation**	**Detection**
Power supplies				
PSH_1	MKBH PS failure under-voltage	Failsafe	No	BEI, BET
PSH_2	MKBH PS failure over-voltage	Failsafe	No	BEI, BET
PSV_1	MKBV PS failure under-voltage	Failsafe	No	BEI, BET
PSV_2	MKBV PS failure over-voltage	Failsafe	No	BEI, BET
Power triggers				
PT_1	Power trigger not responding	Fail silent	Redundancy	Diagnostics
PT_2	Erratic trigger	Failsafe	No	Re-triggering
PT_3	Power supply failure	Failsafe	No	IS
Charging circuits				
CH_1	Capacitor charging failure	Failsafe	No	BEI, BET
CH_2	Capacitor silent failure	Fail silent	No	Diagnostics
CV_1	Resonant circuit charging failure	Failsafe	No	BEI, BET
CV_2	Resonant circuit silent failure	Fail silent	No	Diagnostics
Switches				
SW_1	Failed open	Fail silent	No	Diagnostics
SW_2	Switch erratic	Failsafe	No	BEI, BET
Magnet				
M	Magnet failure	Fail silent	No	Diagnostics
Beam Energy Data Acquisition and Interlocking				
BEA	Error in data acquisition	Failsafe	No	BEI, BETS
BEI_1	Unavailable	Fail silent	No	Diagnostics
BEI_2	False alarm	Failsafe	No	Self announcing
VD	Voltage divider failure	Failsafe	No	BEI, BET

Table 6.6: MKB system failure modes.

to a false alarm. The logic expression is:

$$\begin{aligned} F3_{MSD} = {} & [(DCCT_2 \vee BEA) \wedge \neg(BEI_1 \vee BEI_2)] \vee BEI_2 \vee FC_2 \vee \\ & [(VD \vee \tbinom{15}{1}[M_1]) \wedge \neg(FC_1 \vee FC_2)] \vee \tbinom{75}{1}[TS_2] \wedge \neg(PLC) \end{aligned} \tag{6.13}$$

The FMCCM may also generate false alarms due to high frequency noise on the measured current signal. It is difficult to quantify this contribution and it is considered to be beyond the aim of this study.

6.5 The MKB System Failure Modes

The failure modes of the MKB system are listed in Table 6.6. The failure modes at component level combine together in system failure modes, as listed in Table 6.1. The MKB has silent modes (F1), which can accumulate undetected, and failsafe modes of the F2 and F3 type, which may issue a dump request.

$F1_{MKB}$: **all MKBH or all MKBV magnets are unavailable**. The definition of the failure mode assumes that the loss of either the vertical or the horizontal dilution is a failure concerning safety. The identified causes of this failure mode are: a missed triggering of the power triggers (PT_1), the generators switch not closing (SW_1), the capacitor unable to discharge (CH_2) and the magnet failure (M). The logic expression is:

$$F1_{MKB} = \binom{4}{4}[F1_{MKBH}] \vee \binom{6}{6}[F1_{MKBV}] \tag{6.14}$$

$$F1_{MKBH} = (PT_1 \vee SW_1 \vee CH_2 \vee M) \text{ and } F1_{MKBV} = (PT_1 \vee SW_1 \vee CV_2 \vee M)$$

$F2_{MKB1}$: **The system has failed due to an energy tracking failure**. The logic expression for this failure mode depends on the status of the BEI cards that perform the first level of interlocking before the BETS. If all BEI are available then the energy tracking failure are properly detected whatever the origin. If at least one BEI has failed silent then it is possible that sufficient MKBs have failed to provoke a system failure. Analogously to the MKD system, an over voltage in the generators (PSH_2 and PSV_2) is not tolerated and is not covered by redundancy. This is accounted for in the following expression:

$$\widetilde{F2}_{MKB1} = \dot{\vee}_{k=1\ldots4}[\widetilde{F2}_{MKB1H}(k)] \vee \dot{\vee}_{h=1\ldots6}[\widetilde{F2}_{MKB1V}(h)] \tag{6.15}$$

$$\widetilde{F2}_{MKB1H}(k) = X_{BEIH}(k) \wedge \dot{\vee}_{n=0\ldots k}[\tbinom{k}{n}[F2_{MKB1H}] \wedge \tbinom{(4-n)}{(4-n)}[F1_{MKBH}] \vee \tbinom{k}{1}[PSH_2]]$$

$$\widetilde{F2}_{MKB1V}(h) = X_{BEIV}(h) \wedge \dot{\vee}_{m=0\ldots h}[\tbinom{h}{m}[F2_{MKB1V}] \wedge \tbinom{(6-m)}{(6-m)}[F1_{MKBV}] \vee \tbinom{h}{1}[PSV_2]]$$

$$F2_{MKB1H} = (PSH_1 \vee CH_1 \vee SW_2 \vee PT_2),\ F2_{MKB1V} = (PSV_1 \vee CV_1 \vee SW_2 \vee PT_2)$$

$X_{BEIH/V}(k)$ means that k BEI have failed silent (BEI_1) before the surveyed failure modes (i.e. time priority) and the others are functioning ($\neg(BEI_1 \vee$

$BEI_2)$). The contribution of over-voltage failure mode dominates all other failures. In case the BETS is not available the BEI card is not useful anymore and the failure mode expression (6.15) changes in a simpler expression calculated for k = 4, h =6, without the term $X_{BEIH}(k)$ and $X_{BEIV}(h)$.

The failure mode turns to be safe and detectable by the BETS in case the failure has occurred within the set of the functioning BEI. In this case, the logic expression becomes:

$$\widehat{F2}_{MKB1} = \dot{\vee}_{k=1...4}[\widehat{F2}_{MKB1H}(k)] \vee \dot{\vee}_{h=1...6}[\widehat{F2}_{MKB1V}(h)] \tag{6.16}$$

$$\widehat{F2}_{MKB1H}(k) = X_{BEIH}(k) \wedge \tbinom{(4-k)}{1}[F2_{MKB1H} \vee PSH_2]$$
$$\widehat{F2}_{MKB1V}(h) = X_{BEIV}(h) \wedge \tbinom{(6-h)}{1}[F2_{MKB1V} \vee PSV_2]$$

If the BETS has failed silent the failure mode is not detectable.

$F2_{MKB2}$**: Failure in at least one power trigger power supply**. The logic expression accounts for at least one detected power supplies failures in the power triggers of the MKBH and the MKBV:

$$F2_{MKB2} = \binom{4}{1}[PT_3] \vee \binom{6}{1}[PT_3] \tag{6.17}$$

$F3_{MKB}$**: The internal surveillance has produced a false alarm**. This failure mode includes the failures of the BEA, the BEI false alarm and the voltage divider. The BEA and the voltage divider failures are detected if the BEI is available, that is $\neg(BEI_1 \vee BEI_2)$. The logic expression is:

$$F3_{MKB} = F3_{MKBH} \vee F3_{MKBV} \tag{6.18}$$

$$F3_{MKBH/V} = \tbinom{k}{1}[(VD \vee BEA) \wedge \neg(BEI_1 \vee BEI_2) \vee BEI_2]$$

where k = 4 for the MKBH and k = 6 for the MKBV.

6.6 The Triggering System Failure Modes

The failure modes of the triggering system are listed in Table 6.7. They are combined in the system failure modes of Table 6.1. The triggering system

Failure modes		**Type**	**Compensation**	**Detection**
Client interface				
C_1	Stuck at no dump request	Fail silent	Redundancy	Diagnostics
C_2	Spurious dump request	Failsafe	No	Self-announcing
Oscillator				
O	Loss of RF frequency	Failsafe	No	Surveillance
Phase lock				
PL	Loss of RF phase	Failsafe	No	Surveillance
Dump request storage				
DR_1	Unable to store a dump request	Fail silent	Redundancy	Diagnostics
DR_2	Storage of false dump request	Failsafe	No	Self-announcing
Trigger output gate				
TO_1	The trigger is not transmitted	Fail silent	Redundancy	Diagnostics
TO_2	Spurious trigger	Fail safe	No	Self-announcing
Trigger distribution				
TD	Loss of one TX to power trigger	Fail silent	Redundancy	Diagnostics
Delayed trigger path				
DT_1	Cable/connectors failed	Fail silent	No	Diagnostics
DT_2	Loss of the link to one RTS line	Fail silent	Redundancy	Diagnostics
Clock				
CLK	Clock failure	Failsafe	No	Surveillance
Surveillance				
S_1	Unavailable	Fail silent	No	Diagnostics
S_2	False alarm	Failsafe	No	Self-announcing
VME crate				
VME	Power supplies and fans failure	Failsafe	No	Self-announcing

Table 6.7: Triggering system failure modes.

has fail silent modes (F1) and failsafe modes (F2 and F3 type), which may issue a dump request. As an assumption, all functions implemented in the Field Programmable Gate-Array (FPGA) are assumed to fail independently. The realistic scenario is very difficult to analyse and is not expected to lead to significantly different results.

$F1_{TRG}$**: The system is not able to trigger to an external dump request**. The identified causes are: the failure of the client interface (C_1), the missed storage of the dump request (DR_1), the failure of the trigger output gate (TO_1) in both trigger generators in combination with the failure of the re-triggering path (DT). The logic expression is:

$$F1_{TRG} = F1_{TRG-A} \wedge F1_{TRG-B} \tag{6.19}$$

$$F1_{TRG} = C_{1A/B} \vee [(DR_{1A/B} \vee TO_{1A/B}) \wedge DT]$$

$F2_{TRG1}$: **The system has generated a spurious trigger**. The identified causes of this failure mode are a detected synchronization error (CLK, O, PL), an erratic in the client interface (C_2), in the dump request storage unit (DR_2) or in the trigger output gate (TO_2). Part of this failure is self-announcing and part is failsafe, which depends on the status of the synchronization surveillance X_{SYNC} in branches A and B. The logic expression is:

$$F2_{TRG1} = F2_{TRG1-A} \vee F2_{TRG1-B} \tag{6.20}$$

$$F2_{TRG1-A/B} = C_{2A/B} \vee [(CLK_{A/B} \vee O_{A/B} \vee PL_{A/B}) \wedge X_{SYNCA/B}] \vee DR_{2A/B} \vee TO_{2A/B}$$

The synchronization error is detected if $X_{SYNCA/B} = \neg(S_1 \vee S_2)$.

$F2_{TRG2}$: **The system has generated an asynchronous trigger**. This failure modes occurs if at least one unit among the oscillator, the timing or the internal clock has failed undetected. The logic expression is:

$$F2_{TRG2} = F2_{TRG2-A} \vee F2_{TRG2-B} \tag{6.21}$$

$$F2_{TRG2} = (CLK_{A/B} \vee O_{A/B} \vee PL_{A/B}) \wedge_T X_{SYNCA/B}$$

This failure is undetected if the variable $X_{SYNCA/B} = S_{1A/B}$ and this failure occurred before the surveyed failures (i.e. time priority).

$F3_{TRG}$: **The system has generated a false alarm**. The logic expression for $F3_{TRG}$ is the false alarm of the internal surveillance or detected failures in the VME crate:

$$F3_{TRG} = VME \vee S_{A2} \vee S_{B2} \tag{6.22}$$

The failure of the VME crate concerns the system unavailability only. From the point of view of safety, a possible concern is represented by the internal surveillance that misses the failure of the VME power supplies. This contribution is negligible. Moreover, the triggering system would likely detect this event indirectly as a synchronization error.

Failure modes		**Type**	**Compensation**	**Detection**
Input channel				
IN	One erratic trigger source uncovered	Fail silent	Redundancy	Diagnostics
Output channel				
OUT	Loss of output to one power trigger	Fail silent	Redundancy	Diagnostics
Re-triggering line				
L	Failure of one re-triggering line	Fail silent	Redundancy	Diagnostics
VME	Power supplies and fans failure	Failsafe	No	Self-announcing

Table 6.8: Re-triggering system failure modes.

6.7 The Re-triggering System Failure Modes

The failure modes of the re-triggering system are listed in Table 6.8 and contributes to the the silent failure mode (F1) in Table 6.1. This corresponds to the complete unavailability of the system, which is defined as the loss of both the re-triggering lines A and B:

$$F1_{RTS} = L_A \wedge L_B \tag{6.23}$$

Another risk concerns the loss of the coverage of the erratic trigger at the source where the current is picked up. To demonstrate that this is actually a secondary risk with respect to (6.23), it is necessary to define the re-triggering path. A re-triggering path consists of the input channels that covers the erratic trigger at the source plus the re-triggering lines leading the signal to the re-triggering box connected to the power triggers. A re-triggering action is successfully completed if there are enough re-triggering paths available to re-trigger at least 14 MKD generators. The erratic triggers originated in the power triggers are largely covered as they occur upstream all re-triggering paths. The likelihood of missing one of them is absolutely negligible. The other two sources of erratic triggers (primary and compensation circuits) are covered by two independent input channels, again resulting in a negligible risk of missing the erratic trigger. For this reason, whatever the source of the erratic trigger is, the failure of the re-triggering lines dominates all other possible failure modes within the re-triggering system.

Failure modes		**Type**	**Local detection**	**External detection**
Reference voltage (0,10V)				
REF_1	Incorrect reference voltage	Failsafe	Voting	BETS
MPX analog multiplexer				
AS_1	Multiplexer stuck at an input	Failsafe	Voting	BETS
ADC, analog to digital converter				
ADC_1	Incorrect digital output	Failsafe	Voting	BETS
Averaging module (16 samples)				
AV_1	Incorrect output	Failsafe	Voting	BETS
Transmission				
TX_1	Incorrect transmission	Failsafe	RX surveillance	Voting, BETS
DCCT, current measurement				
$DCCT$	Incorrect output	Failsafe	Voting	BETS

Table 6.9: Beam Energy Acquisition failure modes.

6.8 The BEMS Failure Modes

The failure modes for the BEA and the BEMS are listed in Table 6.9 and Table 6.10 [40]. Only failsafe modes F2 and F3 have been identified[5] and listed in Table 6.1. The 'compensation' column is not useful for these systems and it is replaced by a column including the external detection of the failure mode (if any). $F2_{BEMS}$**: The system delivers the wrong beam energy reference**. This failure is mainly due to errors in the conversion of the measured dipole current to beam energy and its transmission. The logic expression is:

$$F2_{BEMS} = (\binom{4}{1} BEA_{F2} \wedge X_{VT}) \vee BEM_{F2} \tag{6.24}$$

$$BEA_{F2} = (DCCTx \vee DCCTy) \vee (REF_1 \vee AS_1 \vee AV_1 \vee ADC_1) \vee TX_1$$

BEA_{F2} is the failure of the BEA card for which the system delivers the wrong beam energy. The failure is undetected if $X_{VT} = VT_{2A} \wedge VT_{2B}$ and the event occurred before the system failure occurs (i.e. time priority). On the contrary, the failure mode is detected if $X_{VT} = \neg(VT_{1A} \vee VT_{2A}) \vee NOT(VT_{1B} \vee VT_{2B})$.

[5] The failure modes of F1 type, for example those concerning an error in the look-up table for the energy conversion, have been excluded from the analysis because of the rearming procedures that make their likelihood negligible.

Failure modes		**Type**	**Local detection**	**External detection**
RX data receiver				
RX_1	Invalid format (CRC)	Failsafe	RX surveillance	Voting, BETS
RX_2	No data, 1ms timeout	Failsafe	RX surveillance	Voting, BETS
RX_3	Wrong calculated current value	Failsafe	Voting	BETS
RX_4	No value calculated (timeout)	Failsafe	Watchdog	Voting, BETS
RX interlocking				
RXD_1	False alarm	Failsafe	Self-announcing	
RXD_2	Unavailable	Fail silent	Diagnostics	
Voter interlocking				
VT_1	False Alarm	Failsafe	Self-announcing	
VT_2	Stuck at no dump request	Fail silent	Diagnostics	
Watchdog timer				
WDT_1	False Alarm	Failsafe	Self-announcing	
WDT_2	Unavailable	Fail silent	Diagnostics	
Averaging module				
AV_2	Incorrect output	Failsafe	BETS	
AV_3	No value calculated (timeout)	Failsafe	Watchdog	
Current to beam energy reference conversion module				
ER_1	Incorrect output	Failsafe	BEI	
ER_2	No value calculated (timeout)	Failsafe	Watchdog	
Transmission, ANYBUS				
TX_2	Incorrect coding	Failsafe	BETS	
TX_3	Incorrect transmission	Failsafe	BETS	
TX_4	No transmission	Failsafe	BETS	

Table 6.10: BEMS failure modes.

BEM_{F2} is due to a BEM failure for which the incorrect beam energy is delivered. Again, this failure is possible only if the internal surveillance has failed silent before. The expression for the undetected failure is:

$$\widetilde{BEM}_{F2} = [(RX_{A1} \vee RX_{A2}) \wedge_T X_{RXDA}] \vee [(RX_{B1} \vee RX_{B2}) \wedge_T X_{RXDB}] \vee [(RX_{A4} \vee RX_{B4} \vee AV_3 \vee ER_2) \wedge_T X_{WDT}] \wedge_T (X_{VT})$$

where $X_{VT} = (VT_{2A} \wedge VT_{2B})$, $X_{RXD} = RXD_2$, $X_{WDT} = WDT_2$. On the contrary, the expression for the detected failure is:

$$\widehat{BEM}_{F2} = [(RX_{A1} \vee RX_{A2}) \wedge (X_{RXDA} \vee X_{VT})] \vee [(RX_{B1} \vee RX_{B2}) \wedge (X_{RXDB} \vee X_{VT})] \vee [(RX_{A4} \vee RX_{B4} \vee AV_3 \vee ER_2) \wedge (X_{WDT} \vee X_{VT})]$$

where $X_{VT} = \neg(VT_{1A} \vee VT_{2A}) \vee \neg(VT_{1B} \vee VT_{2B})$, $X_{WDT} = \neg(WDT_1 \vee WDT_2)$, $X_{RXDA} = \neg(RXD_{1A} \vee RXD_{2A})$ and $X_{RXDB} = \neg(RXD_{1B} \vee RXD_{2B})$.

$F3_{BEMS}$: **The system has generated a false alarm**. The logic expression for this failure mode accounts for false alarms in the surveillance:

$$F3_{BEMS} = RXD_{A1} \vee RXD_{B1} \vee VT_{A1} \vee VT_{B1} \vee WDT_1 \tag{6.25}$$

6.9 The BETS Failure Modes

The failure modes of the BETS are listed in Table 6.11 [40] and combine together in system failure modes as listed in Table 6.1. The system can fail silent (F1 type) or failsafe (F3 type) with a dump request.

$F1_{BETS}$: **The system is blind to powering failures**. The failure mode accounts for those failure modes that make the system unable to detect an energy tracking error, whatever the source is. This corresponds to the fail silent mode of the BEC module:

$$F1_{BETS} = BEC_1 \tag{6.26}$$

$F3_{BETS}$: **The system has generated a false alarm**. This failure mode accounts for all detected internal failure modes in the BEMS and the BEC[6].

[6] This is classified as a false alarm because the unavailability of a surveillance devices does not directly affect safety.

Failure modes		**Type**	**Local detection**	**External detection**
BEI RX data receiver				
RX_1	Invalid format (CRC)	Failsafe	RX surveillance	Voting, BETS
RX_2	Invalid format from BEAs (CRC)	Failsafe	RX surveillance	Voting
BEI RX interlocking				
RXD_1	False alarm	Failsafe	Self-announcing	
RXD_2	Unavailable	Fail silent	Diagnostics	
BEI averaging				
IC_1	Incorrect output	Failsafe	Voting	
BEI conversion				
ER_3	Incorrect output	Failsafe	Voting	
BEI voter				
VT_1	False Alarm	Failsafe	Self-announcing	
VT_2	No detection, voter failure	Fail silent	Diagnostics	
BEC beam energy controller				
BEC_1	Stuck-at no dump request	Fail silent	Diagnostics	
BEC_2	False alarm	Failsafe	Self announcing	
BEMC for comparison				
$F2_{BEMS}$	Wrong energy reference	Failsafe	IS	BEI
$F3_{BEMS}$	Internal false alarm	Failsafe	Self-announcing	

Table 6.11: BETS failure modes.

The failure of the VME crate that houses all BEI cards is also included.

$$F3_{BETS} = F2_{BEMS} \vee F3_{BEMS} \vee BEC_2 \vee VME \tag{6.27}$$

There are two failure modes of the BEI cards: failed silent (BEI_1) and false alarm (BEI_2). The first failure mode is the silent failure of the voter, namely $BEI_1 = VT_2$. The second failure mode corresponds to any internally detected fault:

$$BEI_2 = [RX_1 \vee RX_2 \wedge [X_{RXD} \vee X_{VT}] \vee [IC_1 \vee ER_3 \wedge X_{VT}] \vee VT_1 \vee RXD_1$$

where $X_{VT} = \neg(VT_1 \vee VT_2)$, and $X_{RXD} = \neg(RXD_1 \vee RXD_2)$.

Chapter 7

Dependability Analysis of the LHC Beam Dumping System

This chapter contains the dependability analysis of the LHC beam dumping system for one year of operation and different operational scenarios.

7.1 From FMECA to Failure Statistics

The logic expression of failure modes, obtained by FMECA in Chapter 6, are transformed into probabilities by applying failure rates statistic at component level. The failure rates are available in manuals, handbooks and manufacturer's datasheets, as average value with 90% confidence level. Formulas exist that adjust them to the operating conditions [66] and apportion to the failure modes that are foreseen for each component [34, 67]. When deriving these figures, some approximations can be applied. For example, the failure rate found in literature for the solid-state switch of the MKD generator is 100 FIT (1 FIT = 10^{-9} failures/h). Taking into account the operational conditions, this is adjusted to 240 FIT and apportioned into 10% failed open mode, 80% failed short mode and 10% slow drift in the electrical parameter. The consequences of slow drifts are negligible due to the post mortem diagnostics, and this quota can be redistributed to the other failure modes

Operator		**Probability**
AND	A $\wedge$ B	$P_A P_B$
OR	A $\vee$ B	$1-(1-P_A)(1-P_B)$
NOT	$\neg$ A	$1-P_A$
XOR	A $\dot{\vee}$ B	P_A+P_B
k out of n	$\binom{n}{k}A$	$\binom{n}{k}P_A^k(1-P_A)^{n-k}$
AND priority	A $\wedge_t$ B	$\int_0^t P_A(t)dP_B(t)$

Table 7.1: Logic operators and probabilities.

according to a conservative approach. Another approximation concerns the statistical independence of failure modes. In the case of the switch the short and open failure modes are mutually exclusive, therefore dependent, see the models described in section 5.1. Nevertheless, the error from assuming them to be independent is small and simplifies the calculations. In the present study, similar approximations have been applied to almost all components.

Once every component failure mode has been attributed a failure rate, the logical expression of the failure modes at system level are transformed into probabilities. The probabilities expressions for the basic logical operators are shown in Table 7.1 for two independent failure modes A and B. For more complex logical expressions, the translation into probabilities may require a pivoting decomposition. The pivoting resolves the ambiguities when a logical term occurs twice or even more times in the expression. Disregarding this aspect of computation would lead to an error because all terms would be assumed independent even when they are not. For example, the logic expression $f(\cdot)$ of the failure mode F is assumed to contain the failure mode A at least twice. To apply pivoting with respect to the failure mode A consists of splitting $f(\cdot)$ into two sub-expressions: $f(A)$ that holds when the A has occurred and $f(\bar{A})$ that holds when A has not occurred. In doing so, the ambiguity is resolved and the two expressions can be translated into probabilities as follows:

$$P[(A \wedge f(A))\dot{\vee}(\neg A \wedge f(\bar{A}))] = P_A P_{f(A)} + (1-P_A)P_{f(\bar{A})} \qquad (7.1)$$

The final result of these mathematical passages is the distribution $P_F(t)$ of the random variable Time To Failure (TTF) of the system failure mode F or its equivalent hazard function $\lambda(t)$ defined in equation (3.1) of section 3.2, where $P_F(t) = 1 - R_F(t)$.

The passages from the FMECA to the calculation of the failure modes statistics are illustrated for the MKD system. The calculations are identical for the other components of the LBDS. The probability expression of the failure mode $F1_{MKD}$, derived from equation (6.1), becomes:

$$P_{F1MKD} = 1 - 15(1 - P_{MKDsilent})^{14} + 14(1 - P_{MKDsilent})^{15} \tag{7.2}$$

$$P_{MKDsilent} = 1-(1-P^2_{PT1})(1-P^2_{SP1})(1-P^2_{SC1})(1-P^2_{CP2})(1-P^2_{COS1-2})(1-P^2_{COS2-2})P_M$$

$$P_{PT1} = 1 - (1 - P^2_{PTM})[1 - \textstyle\int_0^t P_{PTC1}\partial(1 - (1 - P_{PTM-PS})(1 - P_{HV-PS}))](1 - P_{PTC3})$$

$$P_{PTM} = 1 - (1 - P_{PTSP1})(1 - P_{DC1})(1 - P_{PTSC1})(1 - P_{RP1}P_{DP1})$$

The $F2_{MKD1}$ failure mode is detectable or undetectable depending on the status of the 15 BEI cards that survey this failure in the 15 MKD generators, and on the status the BETS, which collects all BEI outputs and decides on dumping the beam, see section 6.3, equation (6.2). The logical expression for the undetected failure mode, given the BETS is functioning, is:

$$\tilde{P}_{F2MKD1} = \sum_{k=1\ldots15} \binom{15}{k} (1 - P_{BEI1} - P_{BEI2})^{15-k} \int_0^t P^k_{BEI1}\partial P(k,t) \tag{7.3}$$

where $P(k,t)$ is a complex expression that describes the probability that an energy tracking related failures may occur in the k uncovered MKD. Another expression is obtained if the BETS has failed silent, see (6.3). In this case the status of the BEI does not enter the formula that becomes:

$$\begin{aligned}\tilde{P}_{F2MKD1nobets} = 1 - [1 - [1 - (1 - P_{MKDenergy})^{15}]14P_{MKDsilent}]\\ [(1 - P_{MKDenergy})^{15} + 15P_{MKDenergy})](1 - P_{PSP2})^{15}\end{aligned} \tag{7.4}$$

The failure mode is safe if the failure has occurred within the set of the functioning BEI. The probability expression for the detectable failure is:

$$\begin{aligned}\hat{P}_{F2MKD1} = \textstyle\sum_{k=0}^{14} \binom{15}{k}(1 - P_{BEI1} - P_{BEI2})^{15-k}P^k_{BEI1}\\ ((1 - P_{MKDenergy})^{15-k}(1 - P_{PSP2})^{15-k})\end{aligned} \tag{7.5}$$

$$P_{MKDenergy} =$$
$$1-(1-P_{PSP1})(1-P_{PSOS1})(1-P_{PSOS2})(1-P_{CP1})^2(1-P_{COS1-1})^2(1-P_{COS2-1})^2$$

On the contrary, if the BETS has failed silent, this failure mode is not detectable and its probability is zero.

The failure probabilities in the above expressions are quantified using the statistics on failure rates at component level, see Appendix A. Results for the MKD failure modes are shown in the plots of Figure 7.1. Where redundancy exists, the failure mode rate starts at zero for t = 0, it reaches a maximum and then tends to a finite asymptotic value, as it is the case of $F1_{MKD}$, see the first plot (top left) of Figure 7.1. Similar behavior is obtained when surveillance and redundancy act together in the same failure mode, as it is the case of $\widetilde{F2}_{MKD1}$, see the third plot of Figure 7.1. The asymptotic value is zero because depends on the reliability of the surveillance, which tends asymptotically to zero. The failure rate decreases with time for all detectable failure modes $\widehat{F2}_{MKD1}, F2_{MKD3}$ and $F3_{MKD}$, as shown in the second, sixth and seventh plot of Figure 7.1. The asymptotic value is zero unless a self-announcing quota exists, which is independent from surveillance. The failure rate of $\widetilde{F2}_{MKD1nobets}$ starts at a certain rate and therefore increases up to a finite value[1]. The failure rate of $F2_{MKD2}$ is constant because all sources of erratic triggers are independent and add up with their constant contribution.

The failure rates for the failure modes of the MKD and the other LBDS subsystems, see Table 6.1, are summarized in Table 7.2. They are given for t = 0, the maximum value (if any), the asymptotic value and the value for t = 10 h, which is an average mission time used in the next section.

[1] A residual non zero failure rate for t = 0 is due to the over voltage failures that are not covered by redundancy.

Failure modes and rates (1/h)			**Rate at T=10 h**	**Rate at T=0 h**	**Maximum**		**Rate Asymptote for t $\rightarrow \infty$**
					Rate	**T [h]**	
$F1_{MKD}$	< 14 MKD available		8.4×10^{-11}	0	1.18×10^{-5}	1.4×10^{6}	8.3×10^{-6}
$F2_{MKD1}$	Energy tracking	~	1.5×10^{-11}	0	6×10^{-8}	57720	0
		~nobets	7.55×10^{-6}	7.5×10^{-6}	-		7.3×10^{-5}
		^	7.7×10^{-5}	7.8×10^{-6}	-		0
$F2_{MKD2}$	Erratic trigger		9.8×10^{-6}	9.8×10^{-6}	9.8×10^{-6}		9.8×10^{-6}
$F2_{MKD3}$	Power supplies failures		3.89×10^{-5}	3.9×10^{-5}	-		0
$F3_{MKD}$	IS fails safely		5.39×10^{-5}	5.4×10^{-5}	-		6×10^{-6}
$F2_{MSD1}$	Energy tracking	~	5×10^{-12}	0	3.7×10^{-8}	2×10^{5}	0
		^	4.99×10^{-6}	5×10^{-6}	-		0
$F2_{MSD2}$	Fast load changes	~	1.3×10^{-11}	0	1.13×10^{-7}	3.2×10^{5}	1×10^{-7}
		^	6.49×10^{-6}	6.5×10^{-6}	-		0
$F3_{MSD}$	IS fails safely		1.05×10^{-4}	1.05×10^{-4}	-		1×10^{-4}
$F1_{MKB}$	No MKBH or no MKBV		4.8×10^{-23}	0	1.4×10^{-6}	1.94×10^{6}	3×10^{-7}
$F2_{MKB1}$	Energy tracking	~	5×10^{-12}	0	4.6×10^{-8}	3.2×10^{5}	0
		~nobets	2.5×10^{-6}	2.5×10^{-6}	8.25×10^{-6}	8.3×10^{5}	5.3×10^{-6}
		^	4.44×10^{-5}	4.45×10^{-5}	-		0
$F2_{MKB2}$	Power supplies failures		1.29×10^{-5}	1.3×10^{-5}	-		0
$F3_{MKB}$	IS fails safely		3.47×10^{-5}	3.48×10^{-5}	-		2×10^{-6}
$F1_{TRG}$	No trigger		2.8×10^{-13}	0	5.5×10^{-8}	4.37×10^{6}	8×10^{-9}
$F2_{TRG1}$	Spurious triggers		8.58×10^{-7}	8.6×10^{-7}	-		0
$F2_{TRG2}$	Synchronization error		4.4×10^{-13}	0	6.35×10^{-8}	4.6×10^{6}	0
$F3_{TRG}$	IS fails safely		2.206×10^{-5}	2.207×10^{-5}	-		2.2×10^{-5}
$F2_{BEMS}$	Energy tracking	~	3×10^{-17}	0	2.7×10^{-8}	3.8×10^{6}	3×10^{-7}
		^	1.19×10^{-5}	1.2×10^{-5}	-		0
$F3_{BEMS}$	IS fails safe		5×10^{-7}	5×10^{-7}	5×10^{-7}		5×10^{-7}
$F1_{BETS}$	Unable to trigger a dump		1×10^{-7}	1×10^{-7}	1×10^{-7}		1×10^{-7}
$F3_{BETS}$	BET fails safe		3.53×10^{-5}	3.5×10^{-5}	-		2.3×10^{-5}
$F1_{RTS}$	Unable to re-trigger		2.2×10^{-12}	0	-		5.19×10^{-7}

Table 7.2: The rates of the LBDS failure modes.

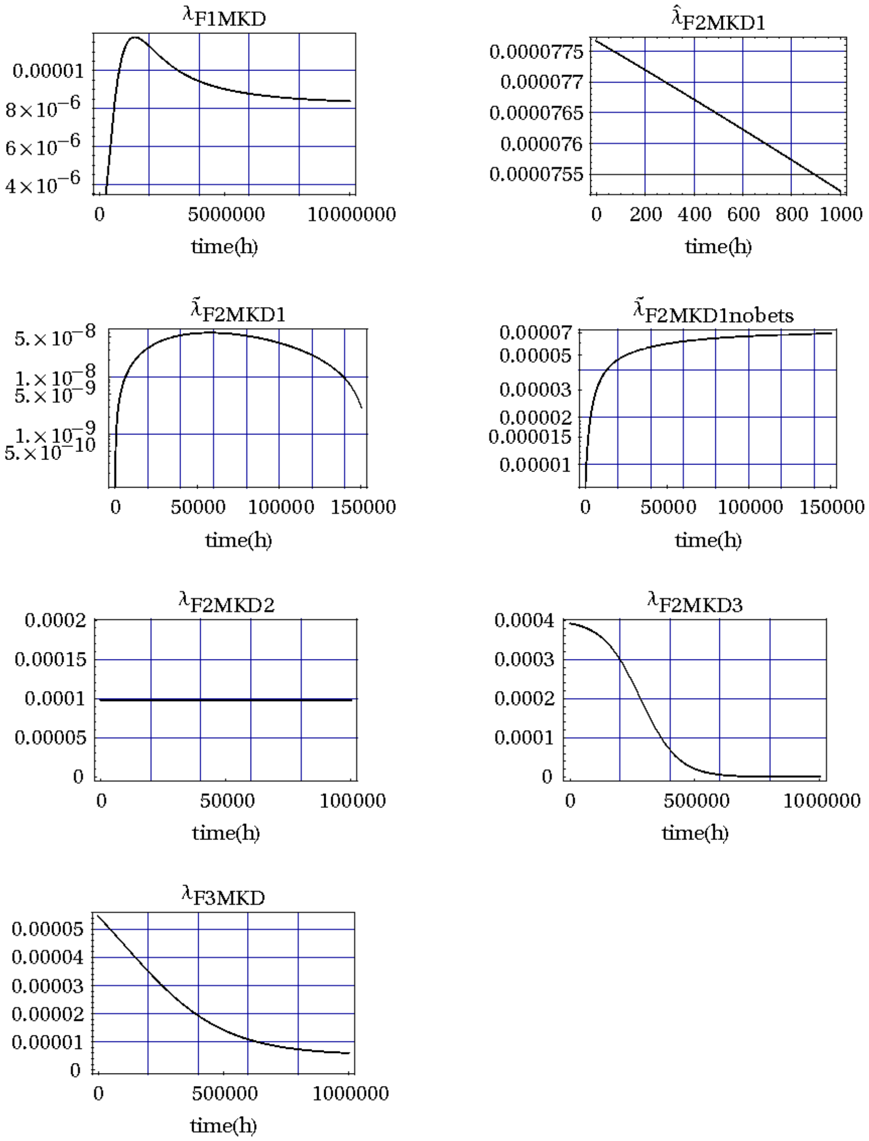


Figure 7.1: The rates of the MKD system failure modes.

7.2 Dependability Modeling of the LBDS

7.2.1 The State Transition Diagrams

A state transition diagram representing the failure process of the LBDS is deduced from the models for not recoverable safety presented in section 5.4, see Figure 7.2. The BETS and the Re-Triggering System (RTS) are explicitly modeled with their failure rates[2]. Six states in total are obtained:

- X0: the system is available.
- X1: the system is available without the BETS.
- X2: the system is available without the RTS.
- X3: the system is available without the BETS and the RTS.
- X4: the system has failed safe.
- X5: the system has failed unsafe.

The model describes the system during the mission time, when the beam is circulating. Once the beam is dumped and the mission is concluded, the system moves into the check phase and post mortem diagnostics is performed. In this phase the system cannot fail and all states are recovered to the initial state X0 with the exception of X5 that is absorbing, see Figure 7.3. Diagnostics and fault recovery aims at restoring full redundancy in the system. If this is the case, all failure rates go back to the value assumed for t = 0, the checks become regeneration points for the failure process and the system is recovered to an 'as good as new state'. In the worst case, if diagnostics is not performed, the redundancy is not restored and the system is said 'as bad as old'. The realistic scenario is supposed to be very close to the as good as new case.

[2] The other Internal Surveillance (IS) is implicitly accounted for in the mechanism that governs the state transitions like in the three state model of subsection 5.4.3.

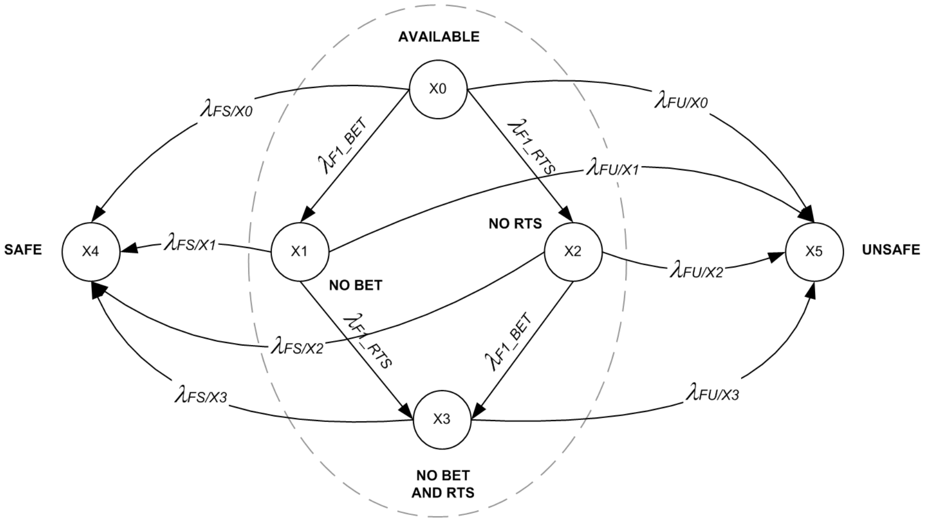


Figure 7.2: The state transition diagram during the mission time.

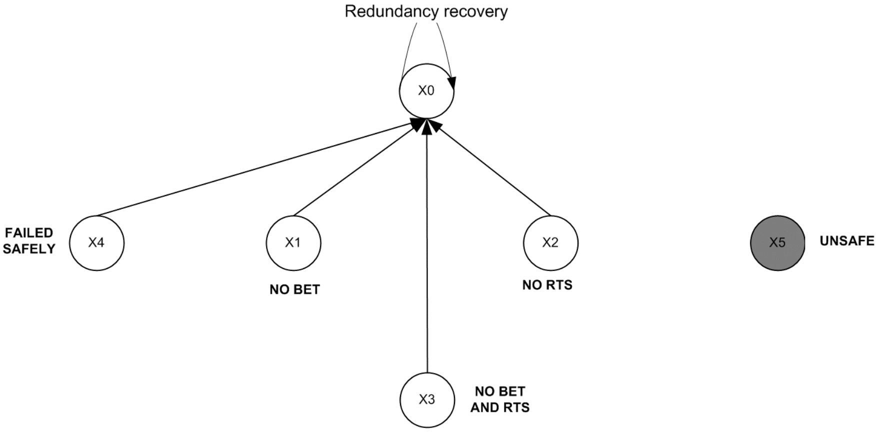


Figure 7.3: The state transition diagram during checks.

Rate	**Expression**
λ_{F1}	$\tilde{\lambda}_{F1MKD} + \tilde{\lambda}_{F1MSD} + \tilde{\lambda}_{F1MKB} + \tilde{\lambda}_{F1TRG}$
$\tilde{\lambda}_{F1MKD}$	$\lambda_{F1MKD} + \tilde{\lambda}_{F2MKD1}(1 - x_1 - x_3)$
$\tilde{\lambda}_{F1MSD}$	$\tilde{\lambda}_{F2MSD1} + \tilde{\lambda}_{F2MSD2}$
$\tilde{\lambda}_{F1MKB}$	$\lambda_{F1MKB} + \tilde{\lambda}_{F2MKB1}(1 - x_1 - x_3)$
$\tilde{\lambda}_{BETS}$	$\tilde{\lambda}_{F2MKD1nobets} + \hat{\lambda}_{F2MSD1} + \tilde{\lambda}_{F2MKB1nobets} + \hat{\lambda}_{F2BEMS}$
λ_{RTS}	λ_{F2MKD2}
λ_{F2}	$\hat{\lambda}_{F2MKD3} + \hat{\lambda}_{F2MSD2} + \hat{\lambda}_{F2MKB2} + \hat{\lambda}_{F2TRG2} + \hat{\lambda}_{F2BEMS} + \lambda_{F3}$
λ_{F3}	$\lambda_{F3MKD} + \lambda_{F3MSD} + \lambda_{F3MKB} + \lambda_{F3TRG} + \lambda_{F3BEMS} + \lambda_{F3BETS}$
$\hat{\lambda}_{BETS}$	$\hat{\lambda}_{F2MKD1} + \hat{\lambda}_{F2MSD1} + \hat{\lambda}_{F2MKB1} + \hat{\lambda}_{F2BEMS}$

Table 7.3: Expressions used in the definition of the state transition rates.

Each state transition in the model of Figure 7.2 is characterized by its sojourn time, which is a random variable with rate $\lambda(t)$, calculated on the basis of the failure rates of Table 7.2. During mission time, the transitions within the available states X0, X1, X2 and X3 are governed by the failures of the BETS and the RTS systems. All other transitions lead outside these four states, either to the unsafe state X5 ($\lambda_{FU/Xk}$) or to the safe state X4 ($\lambda_{FS/Xk}$).

The transition rate $\lambda_{FU/Xk}$ is defined as the rate from the state Xk ($k = 0, 1, 2, 3$) to the state X5. A binary variable x_k is introduced, which is 1 if the system is in the state k and 0 elsewhere. The formula for $\lambda_{FU/Xk}$ is:

$$\lambda_{FU/Xk} = \lambda_{F1} + (x_1 + x_3)\tilde{\lambda}_{BETS} + (x_2 + x_3)\lambda_{RTS} \tag{7.6}$$

where λ_{F1}, $\tilde{\lambda}_{BETS}$ and λ_{RTS} are defined in Table 7.3. The rate λ_{F1} is the common term of the unsafety rate, independent from the state of the system. It accounts for the likelihood that the system has failed unsafe either because of the accumulation of silent failures above a certain threshold, or due to failsafe modes undetected by the internal surveillance and therefore turned to be silent. The other two contributions depend on the status of the BETS and the RTS, which is accounted for by the variables x_k. They give their contribution only if the respective surveillance has failed silent so that the highest unsafe transition rate is found in X3, for which the BETS and the

RTS have both failed silent, see $\lambda_{FU/X3}$ in Figure 7.4. All transition rates leading to the unsafe state X5 are increasing with time[3].

The transition rate $\lambda_{FS/Xk}$ is defined as the rate from Xk ($k = 0, 1, 2, 3$) to the state X4. Again, a binary variable x_k is introduced, which is 1 if the system is in the state k and 0 elsewhere. The formula for $\lambda_{FS/Xk}$ is:

$$\lambda_{FS/Xk} = \lambda_{F2} + (1 - x_2 - x_3)\hat{\lambda}_{BETS} + (1 - x_1 - x_3)\lambda_{RTS} \tag{7.7}$$

where λ_{F2}, $\hat{\lambda}_{BETS}$ and λ_{RTS} are defined in Table 7.3. The rate λ_{F2} is the common term of the failsafe rate similarly to λ_{F1}. It accounts for the system failsafe modes detected by internal surveillance, the self-announcing failures and the internal false alarms. The other two contributions depend on the status of the BETS and the RTS, which is accounted for by the variables x_k. They give their contribution only if the respective surveillance is available so that the highest transition rate is found for X0, for which both BETS and RTS are available, see $\lambda_{FS/X0}$ in Figure 7.5. All transition rate leading to the safe state X4 are decreasing with time[4].

The probability distribution in X = {X0, X1, X2, X3, X4, X5} at time t is the state probability vector $\mathbf{p}(t) = [p_0(t), p_1(t), p_2(t), p_3(t), p_4(t), p_5(t)]$. The calculation of $\mathbf{p}$(t) requires the solution of the Kolgomorov differential equations, see equation (4.10) of section 4.3:

$$\frac{d}{dt}\mathbf{p}(t) = \mathbf{p}(t)\begin{pmatrix} -\lambda_0 & \lambda_{F1BETS} & \lambda_{F1RTS} & 0 & \lambda_{FS/X0} & \lambda_{FU/X0} \\ 0 & -\lambda_1 & 0 & \lambda_{F1RTS} & \lambda_{FS/X1} & \lambda_{FU/X1} \\ 0 & 0 & -\lambda_2 & \lambda_{F1BETS} & \lambda_{FS/X2} & \lambda_{FU/X2} \\ 0 & 0 & 0 & -\lambda_3 & \lambda_{FS/X3} & \lambda_{FU/X3} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{7.8}$$

Due to the failure rates that are not constant, the equation (7.8) describes a Non-Homogeneous Continuous Time Markov Chain (NHCTMC). The (7.8)

[3] This is due to the additional contribution of the failure modes of F2 type remained undetected because of the surveillance failure and turned to be unsafe.

[4] They lose the undetected part that has turned to be unsafe.

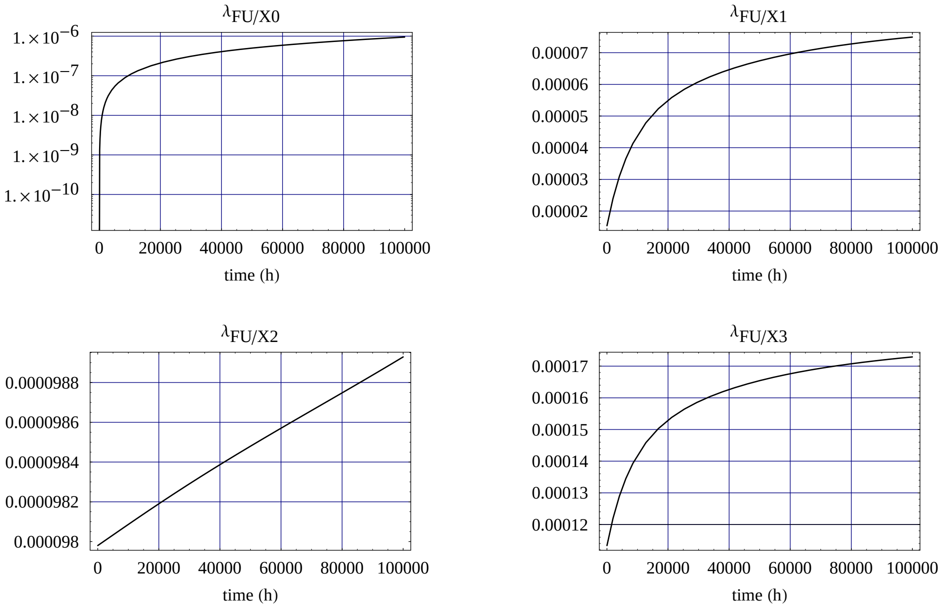


Figure 7.4: Transition rates [1/h] leading to the unsafe state X5.

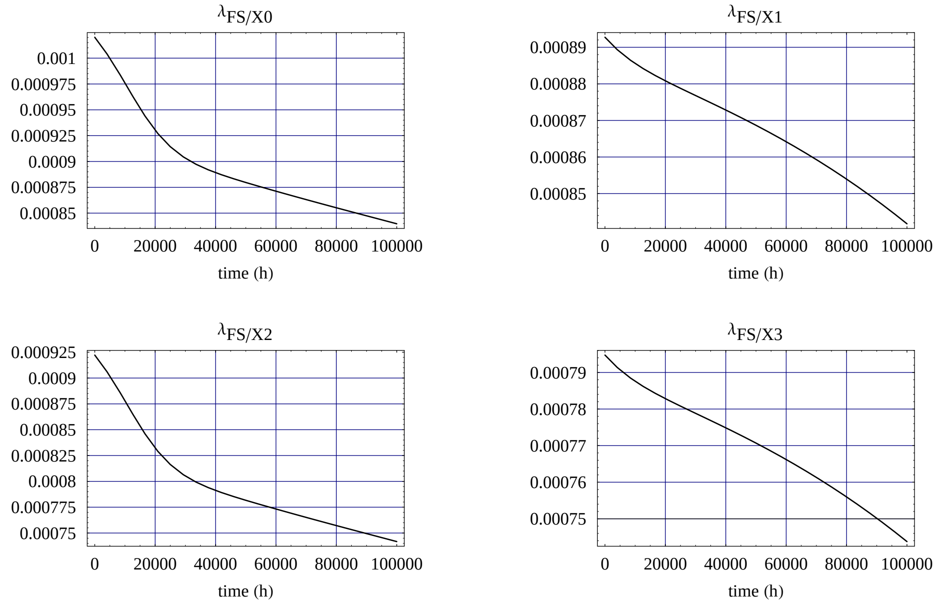


Figure 7.5: Transition rates [1/h] leading to the safe state X4.

is solved for the initial state probability distribution [1,0,0,0,0,0] at t = 0. The diagonal elements of the transition matrix Q are combination of the defined system failure modes, that is:

$$\begin{cases} \lambda_0 = \lambda_{FS/X0} + \lambda_{FU/X0} + \lambda_{F1RTS} + \lambda_{F1BETS} \\ \lambda_1 = \lambda_{FS/X1} + \lambda_{FU/X1} + \lambda_{F1RTS} \\ \lambda_2 = \lambda_{FS/X2} + \lambda_{FU/X2} + \lambda_{F1BETS} \\ \lambda_3 = \lambda_{FS/X3} + \lambda_{FU/X3} \end{cases} \tag{7.9}$$

Safety and availability are the dependability attributes of interest for the LBDS. They are defined in the state space X. Safety $S(t)$ is the probability that the system never moves into the state X5. Availability A(t) is the probability to dump the beam while being in one of the states X0, X1, X2 or X3. In other words, during one mission time, safety is the system availability plus the probability the system has failed safely into X4, that is:

$$\begin{cases} S(t) = 1 - p_5(t) \\ A(t) = S(t) - p_4(t) \end{cases} \tag{7.10}$$

For one year of operation, safety is the probability of never moving into the unsafe state X5. The availability can be related to the number of mission aborts or false beam dumps issued per year, which is a random variable with its probability distribution. The post mortem diagnostics after the beam dump during the check phase is totally irrelevant from the point of view of safety calculation though it may affect the machine availability in a broader sense.

7.2.2 Operational Scenarios

The operational scenario consists of missions and checks that regularly alternate until either the system has failed unsafe or one year of operation is reached. This can be modeled as a timed stochastic Petri net [25], see Figure 7.6 and Table 7.4. One token marks the states of the net by moving in five places {MISSION, CONTINUE, CHECK, STOP, COUNTER}. The system

Transition	Type	Guard
T_{DR}	Random	-
T_C	Random	-
T_{FD}	Instantaneous	X = X4
T_{No}	Instantaneous	(X = X5) OR 1 year
T_{Yes}	Instantaneous	X = {X0,X1,X2,X3}

Table 7.4: The transitions of the Petri net.

waits in the place MISSION for the firing of the random transition T_{DR}, corresponding to an external dump request, or the instantaneous transition T_{FD}, that fires when the status X of the system moves to X4. After the beam dump, the token moves into the place CONTINUE. Two instantaneous transitions can be enabled depending on the status X of the system. T_{No} is enabled if X = X5 or one year of operation is reached[5] and the token moves to STOP. T_{Yes} is enabled if X = {X0, X1, X2, X3} and the token moves to CHECK. The token also moves to the CHECK place in case of false dump. After leaving the CHECK place, one token is deposited into COUNTER and another moves to MISSION for the start of a new operation. The resulting marking of the Petri net identifies the system when this is running a mission or performing checks. For example, the marking $< 1, 0, 0, 0, k >$ means that the system is running the k+1 mission after k safe missions. The Petri net and the subordinated failure process in X are mutually dependent, as attested by the guards that enable the transitions of the Petri net depending on X, and the dynamic of the failure process that changes depending on the phase executed. The combination of the Petri net and the subordinated processes results in the state transition diagram of Figure 7.7. The macro state XAk is the set of available states X0, X1, X2 and X3 for the mission k.

Three different operational scenarios (OP) are assumed for the dependability analysis of the LBDS. They depend on the characterization of the

[5] This event can be related to the number of performed missions in COUNTER or to a variable T, i.e. the cumulated operational time, which can be represented as a reward function in the Petri net.

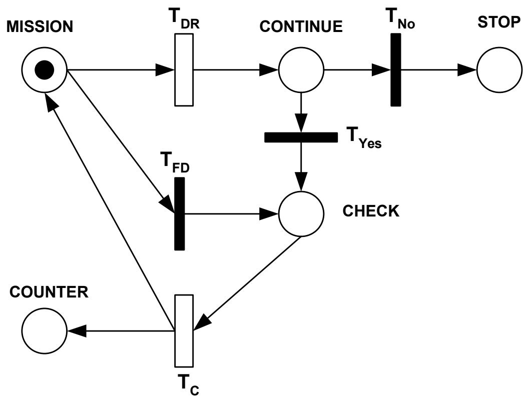


Figure 7.6: A Petri net for the representation of missions and checks.

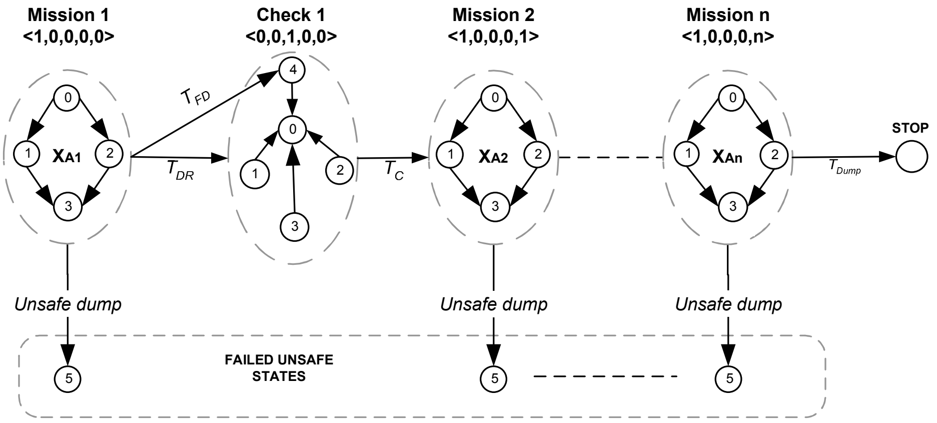


Figure 7.7: The state transition diagram generated from the Petri net.

Operational scenario	Mission length	Regeneration at check
OP1	Deterministic	Yes
OP2	Deterministic	No
OP3	Random	Yes

Table 7.5: The operational scenarios.

mission time (deterministic vs. random) and the check (regeneration points vs. no regeneration points), see Table 7.5. For the three operational scenarios considered, an analytical description exist. OP1 and OP3 describe stochastic Markov regenerative processes (the OP1 is by far simpler to analyse than OP3), see section 4.4. In OP2 the solution can be obtained by reconfiguring the Markov chain (i.e. transition matrix and initial state probability) at every new mission[6]. The equation for the description of the various models have been solved numerically within the *Mathematica*® framework [94].

7.3 The Operational Scenario 1

7.3.1 Dependability Analysis

The dependability analysis of the first operational scenario is based on the following assumptions:

- (A1) 400 missions, 10 hours each, alternate to 2h hours of checks for 200 days of operation.
- (A2) The system is recovered to an 'as good as new' state after each check.
- (A3) Operation is stopped if the system has failed unsafe.

The assumption A1 defines a deterministic operational scenario. The 10 hours mission length accounts for an average exploitation of the machine,

[6] The case study without regeneration points and random mission can only be treated by Monte-Carlo simulation, which also applies to the other three cases. This last scenario is not taken into consideration.

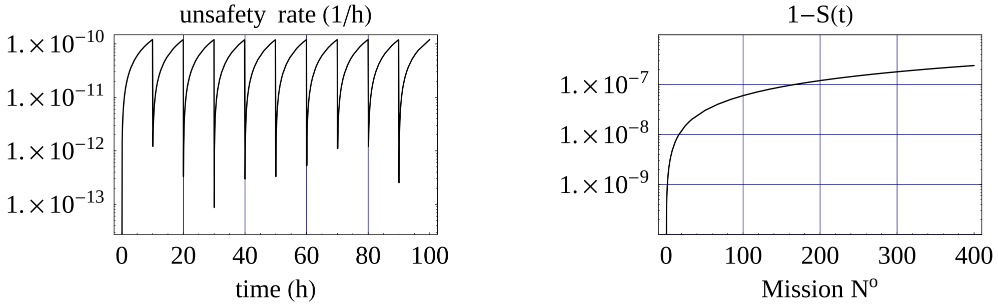


Figure 7.8: The unsafety rate (left) and the unsafety of the LBDS (right).

while the 2 hours check accounts for a reasonable interval including the post mortem diagnostics (up to a few minutes) and the rearming procedures. The assumption A2 states that the failure processes are repeated identically for each mission. As a consequence, it is possible to solve the Kolgomorov equations (7.8) for one mission time instead of for 400 missions, and arrange the result like in equation (5.38) of section 5.4.2. The formula used to calculate the unsafety $U(t) = 1 - S(t)$ for one year of operation is:

$$\begin{cases} U(t) = 1 - S(\tau)S(10)^{k-1} & k > 0 \\ U(t) = 1 - S(t) & k = 0 \end{cases} \tag{7.11}$$

where $t = 10k+\tau$ and $0 \leq \tau \leq 10$. The time between two operational periods does not enter the calculation, as during this period the system cannot fail. The unsafety $U(t)$ versus the number of mission is shown in Figure 7.8 (right). It starts at zero for $k = 0$, and increases to 2.418×10^{-7} for $k = 400$, which is one year of operation. The MTTUF, calculated in equation (5.39), is 1.65×10^{10} hours and corresponds to a rate of 6.05×10^{-11}/h that is largely SIL4. The unsafety rate or residual hazard rate is shown in Figure 7.8 (left) for 10 consecutive missions, check downtimes excluded. The effect of diagnostics results in the regeneration of the unsafety rate after each mission, which becomes periodic.

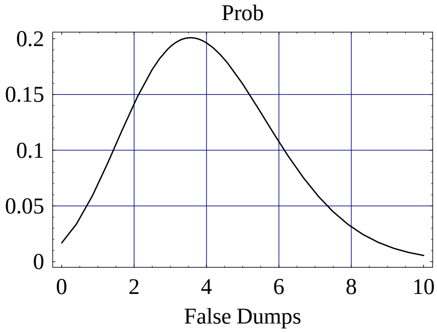


Figure 7.9: Probability distribution of the number of false dumps for one year of operation.

The expected number of false dumps N is Binomially distributed with parameter $p = p_4(t)$ for $t = 10$ h and 400 missions:

$$P[N = n] = \binom{400}{n} p^n (1-p)^{400-n} \tag{7.12}$$

The resulting distribution is shown in Figure 7.9. The average value and the standard deviation are respectively $400 \times p$ and $\sqrt{[400(1-p)p]}$, which results in 4.06 false dumps per year, with ±2.03 as standard deviation (80% confidence level).

7.3.2 Criticality Analysis

The system unsafety can be apportioned to the various subsystems in order to identify those that take the largest contribution. The unsafety is rewritten as $U(t) = U_0(t)+U_1(t)+U_2(t)+U_3(t)$, where $U_k(t)$ is the probability of failing unsafe from one of the four states X0, X1, X2 and X3. For one year of operation, $U_0 = 2.147 \times 10^{-7}$, $U_1 = 0.271 \times 10^{-7}$ and $U_2 = 0.88 \times 10^{-11}$. U_3 is negligible being several order of magnitude smaller than U_2. Each U_k $(k = 0, 1, 2, 3)$ is apportioned to the LBDS sub-systems. The probability that a subsystem failure is at the origin of the failure of the system is an a-posteriori conditional probability and can be calculated in good approximation using the average of the failure rates over a mission time interval. For example, the conditional probability that the MKD system is the cause of the transition to X5 from X0 is:

$$\Pi^0_{MKD} = (\Lambda_{F1MKD} + \Lambda_{F2MKD1})/\Lambda_{FU/X0} \tag{7.13}$$

where $\Lambda_Y = \frac{1}{10}\int_0^{10} \lambda_Y dt$ is the average failure rate over one mission time interval. The fraction of $U(t)$ caused by the MKD system is:

$$U^{MKD} = \sum_{k=0,1,2,3} \Pi^k_{MKD} U_k \tag{7.14}$$

so that the MKD system contributes to the unsafe failure of the LBDS with the percent $\rho_{MKD} = U^{MKD}/U$. Identical calculations are made for the other subsystems.

The resulting apportionment of unsafety to the LBDS components is shown in Table 7.6. In total, 99.5% of unsafety is due to the complete magnets assembly (i.e. magnets plus power converters) with the remaining 0.5% due to the triggering and the beam energy tracking electronics. The largest contribution is due to the MKD system, which is responsible for 75% of unsafety. The analysis can be continued by investigating the criticalities inside each sub-system. The MKD bottlenecks for safety are the magnet failure, which is masked by 14 out of 15 redundancy, and the over-voltage of the power supplies in the generator, which is surveyed but not masked by redundancy. The MSD bottleneck is the short-circuit in the magnet coil or in the power converter, which leads to a fast magnet current change. The MKB bottleneck is the over-voltage of its generators, which is monitored but not masked. The trigger output gate is the bottleneck of the triggering system, which is masked by 1 out of 2 redundancy. The critical part of the BEMS is the averaging of the dipole currents, the beam energy conversion and the transmission to the MKD and MKB generators, which remain totally uncovered as soon as the BETS has failed silent.

In a similar way, the false dumps can be apportioned to the LBDS subsystems. Results are shown in Table 7.7. The expected biggest source of false dumps is again the MKD system (61%), followed by the MKB (20%),

System	**From X0**	**From X1**	**From X2**	**Total%**
MKD	78.14	48.8	$\cong 1$	74.8
MSD	16.9	32.4	9.8×10^{-6}	18.6
MKB	4.7	16.3	2.5×10^{-6}	6.1
BEMS	0	2.5	0	0.27
Triggering	0.26	9.0×10^{-7}	1.4×10^{-7}	0.23

Table 7.6: Apportionment of unsafety to the LBDS components.

MKD	**MSD**	**MKB**	**Triggering**	**BEMS**	**BETS**
2.47 (61.4%)	0.46 (11.4%)	0.83 (20.6%)	0.08 (2.16%)	0.05 (1.25%)	0.14 (3.5%)

Table 7.7: Apportionment of false dumps to the LBDS components.

the MSD (11%) and the electronics (triggering system, BEMS and BETS). At component level, the largest contribution to false dumps is expected from the failure of the power trigger power supplies with about 2 per year. A negligible contribution is expected from the fast current change in the MSD, with 0.02 per year, and the detected synchronization failures in the triggering system. The false dumps can also be apportioned with respect to the failure modes. Four types of failures are considered: the energy tracking failures, the erratic triggers, the failure modes detected by Internal Surveillance (IS) and the false alarms. The results are shown in Table 7.8. Both energy tracking failure and erratic triggers are expected to give a minor contribution to the false dumps. The erratic triggers also represent the expected number of asynchronous dumps per year, $\cong 0.4$ in total. A significant fraction of false dumps is calculated to come from false alarms in the surveillance systems, about 1 per year.

Energy tracking	**Erratic triggers**	**Others IS**	**False alarms**
0.55 (13.6%)	0.39 (9.6%)	2.1 (51.95%)	1 (24,7%)

Table 7.8: Apportionment of false dumps to the failure modes.

7.3.3 Sensitivity Analysis

The aim of sensitivity analysis is to quantify the effect on the system dependability if some of the fault tolerant design features were totally or partially removed, provided the functionality is not compromised. It also investigates the sensitivity of the calculated safety and number of false dumps with respect to the model parameters, i.e. the component failure rates. Through sensitivity analysis it is also possible to discover if safety and availability, which are a design trade-off, are well balanced in the system.

Sensitivity to the Fault Tolerant Design

Redundancy exist at various levels in the MKD system, see Figure 2.3. At high level 14 out of 15 redundancy permits to withstand the failure of one MKD assembly (i.e. magnet and/or generator). The removal of one MKD assembly would results in an unsafety of 0.011 per year, namely 2.75×10^{-6}/h that is SIL1, for a number of false dumps of 3.89 on average per year. At lower level, one generator branch can be removed inside each MKD generator without compromising its functionality. The resulting unsafety is 2.34×10^{-6} per year, which is still SIL4, while the number of false dumps decreases[7] to 3 per year. This result demonstrates that redundancy is more important if applied at system level (i.e. 14 out of 15) than at sub-system or component level. A further demonstration is given for the power triggers of the MKD and MKB kickers, see section 2.2.1. The removal of one the two modules working in parallel in the power trigger is less important because it is embedded in 1 out of 2 generator branches within 14 out of 15 redundancy of the MKD system. The unsafety remains practically unchanged, 2.420×10^{-7} per year, with a saving of 0.2 false dumps per year. Redundancy can also be removed from the triggering system using one instead of the two trigger generators in parallel, see Figure 2.8. In this case the unsafety increases to about 4.7×10^{-4}

[7] The saved fraction of false dumps is less significant if considered in the global context of the machine protection system, see Chapter 8.

Case studied	**Unsafety/year**	**False dumps/year**
Default scenario	2.41×10^{-7} (> SIL4)	4.06
No redundant power triggers	2.34×10^{-6} (SIL4)	3.02
No redundant triggering sys.	4.68×10^{-4} (SIL2)	4.02
14 MKD	0.011 (SIL1)	3.89
No BETS	0.059 (< SIL1)	3.40
No RTS	0.32 (< SIL1)	4.06

Table 7.9: Sensitivity to fault tolerant design and surveillance.

per year, which is 1.17×10^{-7} per hour. This is SIL2 and is not acceptable[8].

The results are also sensitive to on-line surveillance of the energy tracking failures in all magnet power converters and the erratic triggers in the MKD. The removal of the BETS and the BEI-BEA cards, used for analysing the signals from the magnet power converters, leaves the system uncovered with respect to any energy tracking failure. Only the BEMS is kept for the beam energy measurement and distribution. This scenario is described by equation (7.8) for an initial state probability distribution [0,1,0,0,0,0] corresponding to the state X1. The unsafety would increase to 0.059 per year, namely 1.4×10^{-5}/h, which is not even SIL1. If the re-triggering system is removed, any erratic trigger in the MKD becomes unsafe. This scenario is described by equation (7.8) for an initial state probability distribution [0,0,1,0,0,0] corresponding to the state X2. In this case, unsafety dramatically increases to 0.32 per year that corresponds to 8×10^{-5}/h as equivalent failure rate (< SIL1). Results of the presented sensitivity analyses are summarized in Table 7.9.

Sensitivity to the Failure Rates

The calculated system dependability is sensitive to the variation of the parameters of the model, i.e. the failure rates. The LBDS is designed in order to have no single point of failure. For this reason, the variation of the failure

[8] The triggering system is upstream the MKD and does not benefit of the 14 out of 15 redundancy.

Failure rate	**Safety**	**False dumps/year**
MKD magnets ×100	SIL2	4.06
MKD power supplies ×100	SIL4	25
BEI cards ×100	SIL4	24
Power trigger module ×10	>SIL4	16.5

Table 7.10: Sensitivity to the failure rates.

rates is expected to be less important for safety, while a significant increase of the number of false dumps is expected at detriment of availability. Some examples are given for the magnets, the power supplies of the MKD system, the BETS data acquisition channels and the power trigger modules, see Table 7.10. One MKD magnet failure is masked within 14 out of 15 redundancy. For an assumed two order of magnitude larger magnet failure rate λ_m the system unsafety would increases from 2.418×10^{-7} to 0.0013 per year, which is SIL2, crossing SIL3 around $60\lambda_m$. The MKD is sensitive to the failure rate of the power supplies in the MKD generators, 45 in total, in particular to the over-voltage failure mode that is not tolerated by 14 out of 15 redundancy. If the assumed default value of 1×10^{-6}/h increases of two order of magnitude the safety remains still SIL4, thanks to the BETS surveillance that detects the failure. As a drawback, these detected failures add up to the number of false dumps, which increases from 4 to 25 on average per year. Similar results are obtained for the failure rates of the BEI cards. If these increase of two orders of magnitude, unsafety would be around 8×10^{-6} per year, which is still SIL4, for 24 false dumps per year, see Table 7.10. Another very interesting example concerns the power trigger modules of the power trigger system. The redundancy of two trigger modules in parallel per power trigger takes a negligible contribution to safety. If the failure rates in the power trigger modules is increased one order of magnitude the safety remains practically the same 2.36×10^{-7}, for a number of false dumps that becomes 16.5 per year[9]. This is a non acceptable trade-off and shows a potential weakness in

[9] As a collateral effect of the augmented number of false dumps, the mission time is shortened, and the system has less time to fail unsafe, which explains the better value

the design of the power trigger from the availability point of view.

7.4 The Operational Scenario 2

7.4.1 Dependability Analysis

The analysis of the second operational scenario is based on the same assumptions A1 and A3 given for the operational scenario 1, plus a new assumption A2:

- (A2) The check is performed only if the system has failed safely.

The new assumption A2 says that the system is 'as bad as old' at the start of every new mission for those failures that are masked by redundancy and may accumulate undetected. Only if the surveillance has detected the failure and issued a false dumps, the check is performed and the system is recovered to an 'as good as new' state. The resulting stochastic process is not Markov regenerative and its solution must be calculated for one year of operation instead of one mission time like in OP1. The transition matrix Q and the initial conditions in the Kolgomorov equations (7.8) are recalculated at every new mission k according to the following formulas:

$$\begin{cases} \mathbf{Q}(1,t) = \mathbf{Q}(t) \\ \mathbf{Q}(2,t) = p_4(1,T)\mathbf{Q}(1,t) + [1 - p_4(1,T) - p_5(1,T)]\mathbf{Q}(1,t+T) \\ \cdots \\ \mathbf{Q}(k+1,t) = p_4(1,T)\mathbf{Q}(1,t) + [1 - p_4(k,T) - p_5(k,T)]\mathbf{Q}(k,t+T) \end{cases} \tag{7.15}$$

where $0 \leq t < 10$ and $k < 400$. The initial conditions are:

$$\begin{cases} \mathbf{p}(1,0) = [1,0,0,0,0,0] \\ \cdots \\ \mathbf{p}(k+1,0) = [p_0(k,T) + p_4(k,T), p_1(k,T), p_2(k,T), p_3(k,T), 0, p_5(k,T)] \end{cases} \tag{7.16}$$

obtained for safety.

The matrix $\mathbf{Q}(k+1,t)$ is the transition matrix in the mission k+1. It consists of two terms. The first term is the 'recovered part' $p_4(1,T)\mathbf{Q}(1,t)$ which is the probability that the system has failed safely in the previous mission k multiplied by the transition rate matrix regenerated at t = 0. The other term is the 'aging part' $[1 - p_4(k,T) - p_5(k,T)]\mathbf{Q}(k,t+T)$ which is the availability of the system in the mission k multiplied by the transition rate matrix, updated for a time interval T = 10 h that accounts for the missed regeneration. In a similar way, the initial state probability vector is recovered to X0 only for the fraction of false dumps $p_0(k,T)+p_4(k,T)$. The algorithm used for the calculation of safety for one year of operation is the following:

```
< Initialize Safety array of length 400 >
< Initialize False dumps array of length 400 >
...
/* Loop for missions 1 to 400 */
For}k = 1 to 400
{
/* Update transition rates matrix */
switch (k)
{
     case k==1:
     {Q(1,t) = Q(t); p0=[1,0,0,0,0,0]};
break;
default:
{Q(k,t) = False_dumps[k-1]Q(0,t)+(Safety[k-1]-False_dumps[k-1])]Q(k-1,t+T);
p0 = [p_1(k,T)+p_1(k,T),p_1(k,T),p_2(k,T),p_3(k,T),0,p_5(k,T)]};
break;
}
p(k,t) = Markov[Q(k,t), p0];
Safety [k] = (1 - p_{5}(k,T));
False_dumps [k] = p_{4}(k,T);
};
```

The function 'Markov' returns the solution of (7.8) for the transition matrix $\mathbf{Q}(k,t)$ and the initial distribution $\mathbf{p}(k,0)$ as calculated by the equations (7.15) and (7.16).

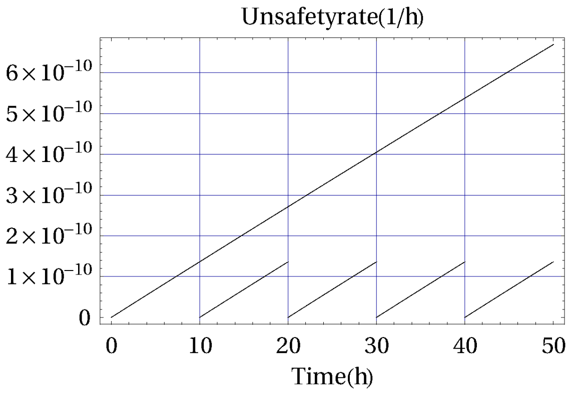


Figure 7.10: Comparison of the unsafety rate for the operational scenario 1 and the operational scenario 2.

The average number of false dumps is still about 4.0 for 400 missions, which means that the system is recovered ‘as good as new’ four times per year, corresponding to a regeneration interval of 80 missions on average. The resulting unsafety is 3.15×10^{-5} (SIL4), two orders of magnitude larger than the unsafety calculated for the previous operational scenario. This can be explained by comparing the unsafety rate for the operational scenario 1 and for the operational scenario 2. They are shown in Figure 7.10 (left) for five consecutive missions. The unsafety rate for the operational scenario 2 keeps increasing except for a little percentage, not appreciable in the plot, which is regenerated at the false dump.

The operational scenario 2 could be further developed for quantifying the benefit of a maintenance program that includes inspections policies for the various parts of the system. This issue is not treated in the present study.

7.5 The Operational Scenario 3

7.5.1 The Markov Regenerative Model

The analysis of this third operational scenario is based on the assumptions A2 and A3 of the operational scenario 1, plus a new assumption A1:

- (A1) Missions of random duration alternate to checks of 2 hours for 200 days (i.e. 4800 hours) of operation.

The assumption A1 describe a random operational scenario. The mission length is determined by the occurrence of a dump request, which can be internal like a false dump (FD), or external (DR) like a planned dump request or other dump requests (e.g. beam induced, machine protection). These events are the regeneration points of the Markov regenerative process (MRGP) that governs the status of the system over one year of operation (see section 4.4). The subordinated process o the MRGP is the Markov chain described by equation (7.8), provided that the rates of the external dump requests are added to all transition rates[10] leading to the state X4. The embedded renewal sequence is governed by the dump request events and contains the states {X0k}, for k being the number of mission. Safety is calculated by the **generalized Markov renewal equation** defined in (5.26) of section 5.4:

$$S(t) = A(t) + \int_0^t S(t-\tau)dp_4(\tau) \tag{7.17}$$

where $S(0) = A(0) = 1$. $A(t) = p_0(t) + p_1(t) + p_2(t) + p_3(t)$ and $p_4(t)$ are calculated using (7.8). The solution of (7.17) is achieved numerically using as reference value $S(t) = (\hat{S}(t) + \check{S}(t))/2$, with $\hat{S}(t)$ and $\check{S}(t)$ calculated by (5.29) and (5.30) respectively.

The number of missions performed over one year is a **renewal counting process** (see section 4.2). The average number of missions per year is derived from equation (5.42) of section 5.4.4:

$$M(t) = p_4(t) + \int_0^t M(t-\tau)dp_4(\tau) \tag{7.18}$$

[10] The rates can be added as they are statistically independent events.

where $M(0) = 0$. The solution is obtained numerically using as reference value $M(t) = (\hat{M}(t) + \check{M}(t))/2$ with $\hat{M}(t)$ and $\check{M}(t)$ calculated by (5.43) and (5.44) respectively.

The results obtained from the equations (7.17) and (7.18) are the average values for safety and the number of missions for one year of operation. The figures must be adjusted by including the 2h checks downtime. The following steps are applied to the analysis of the operational scenario 3:

- **Step 1**: calculate the safety by solving (7.17) for t = 4800 h:

  $S(t) = \frac{1}{2}(\hat{S}(t) + \check{S}(t))$

- **Step 2**: calculate the number of performed missions and the average mission time T_M by solving (7.18) for t = 4800 h:

  $T_M = 4800/M(4800)$ where $M(t) = \frac{1}{2}(\hat{M}(t) + \check{M}(t))$

- **Step 3**: calculate the expected number of missions per year M_{year} including the 2 hours check in between missions:

  $M_{year} = 4800/(T_M + 2)$

- **Step 4**: recalculate the safety for one year of operation now including the 2 hours check:

  $S_{year} = S(T_M \times M_{year})$

- **Step 5**: calculate the number of false dumps per year.

The steps 1 - 4 adjust the results to the operational profile. The last step derives the number of false dumps from the number of missions performed. A way to do this is to split the state X4 into two states: one $X4_{FD}$ for the false dumps, and the other $X4_{DR}$ for the external dump. The Markov chain is solved for one mission length and provides the probability $p_{4FD}(t)$ of being in the state $X4_{FD}$, $p_{4DR}(t)$ of being in the state $X4_{DR}$. The probability that the mission is terminated by a false dump is $\frac{p_{4FD}(t)}{p_{4FD}(t)+p_{4DR}(t)}$, which depends on t. As average value is taken the following quantity:

$$\rho_{FD} = \lim_{T\to\infty} \int_0^T \frac{1}{T} \frac{p_{4FD}(t)}{(p_{4FD}(t) + p_{4DR}(t))} dt \tag{7.19}$$

The same formula can be used for the apportionment of the others dump requests with their distributions.

Two scenarios are analysed for OP3: OP3A that includes planned dump request of random duration and OP3B that includes the same planned dump request of OP3A plus the dump request caused by the beam instabilities during the early period of LHC operation.

7.5.2 Analysis of the Operational Scenario 3A

The planned dump requests are modeled with a Weibull[5,11] distribution, see equation (3.2), which corresponds to 10.1 hours average and ±2.31 standard deviation.

The calculated lower and upper bound for unsafety are $\hat{S} = 2.9507 \times 10^{-7}$ and $\check{S} = 3.259 \times 10^{-7}$ at t = 4800 h, for a mean value of 3.105×10^{-7}, see Figure 7.11. The number of missions at t = 4800 h is within 455 and 503, for a mean value of 479 missions with a 10.02 hours average mission time. This value is smaller than the assumed 10.1 average length of the planned dump requests because of the additional false dump requests. The results are adjusted following steps 3 and 4. The number of missions M_{year} that alternate with 2 h check is recalculated and results in 399 per year on average. The unsafety ranges between $\hat{S} = 2.452 \times 10^{-7}$ and $\check{S} = 2.708 \times 10^{-7}$ for a mean value of 2.58×10^{-7} that is very close to 2.418×10^{-7} obtained for the operational scenario 1. The number of false dumps is calculated using (7.19) and it is 4.1 per year, which is practically the same value that was obtained for the operational scenario 1.

The Accuracy of the Numerical Solution

The accuracy of the numerical solutions for safety and for the number of missions is analysed for a 240 hours time interval versus the integration step, taken between 1/8 and 8 hours[11]. The accuracy is the difference between

[11] The demonstration does not depend on the chosen time interval length.

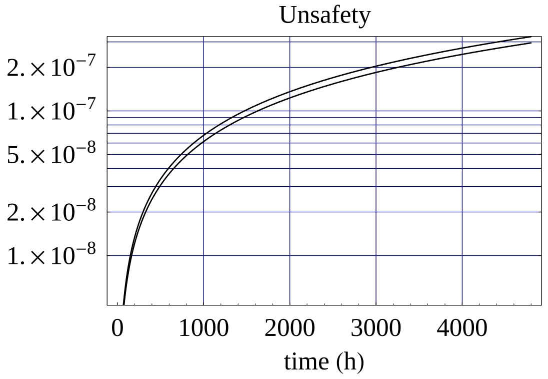


Figure 7.11: Unsafety upper and lower bound solutions for one year of operation, assuming a Weibull[5,11] for the planned dump request.

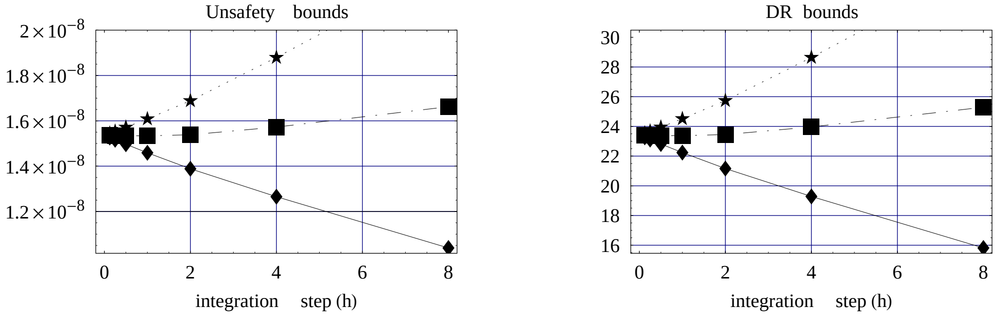


Figure 7.12: Upper bound, lower bound and mean value after 240 hours of operations, versus the integration step, for unsafety (left) and the number of dump requests DR (right).

the upper and lower bound solutions, that is (5.29) and (5.30) for safety, (5.43) and (5.44) for the number of missions. This difference reduces as the integration step increases, see figure 7.12. In addition, the arithmetic mean of safety seems insensitive to the integration step up to about 2 h where it starts diverging as shown in Figure 7.12 (left). The same holds for the mean of the number of missions, as shown in Figure 7.12 (right). This result confirms the possibility to use the arithmetic mean as an estimator of $S(t)$ with the taken integration step of 1 hour[12].

[12] The solution algorithm is $O(2^n)$, where n = 4800/integration step.

Sensitivity to the Dump Request Distribution

The sensitivity to the planned dump requests Weibull distribution is analysed with respect to the variance and the average. For the variance, three Probability Density Functions (PDF), Weibull[3,11.3](A), Weibull[5,11](B) and Weibull[10,10.61](C) with the same average of 10 hours and different variance ±3.57, ±2.31 and ±1.25 hours respectively, are shown in Figure 7.13. The result for unsafety, false dumps (FD), number of missions (M_{year}) and average mission time (T_M) are listed in Table 7.11. A larger variance turns in more unsafety for the system, which is explained by the increasing fail unsafe rates, see Figure 7.4.

The results are also sensitive to the average value of the Weibull distribution. Two PDF, Weibull[5,7](D) and Weibull[10,15](E) with averages 6.4 and 13.8 hours respectively are shown in Figure 7.13. The results are shown in Table 7.11. Again, the differences between figures can be explained by the increasing failure rate. The probability of failing unsafe is higher for a long average mission time than for a short one. Similar considerations hold for the false dumps for which a long mission has a higher probability of being aborted than a short one. In general, a longer mission time implies a lower safety but a higher exploitation of the machine. For example, the Weibull[5,11] has 399 missions 10.02h each on average that makes 3998 hours of LHC operation with the beam, while the Weibull[5,7] yields 573 shorter missions 6.37h each, which makes only 3650 hours of LHC operation with the beam per year. The impact of the 2 hours downtime for checks is determinant. It is also true that shorter missions have the advantage of a higher beam luminosity. This is a trade-off between safety and the machine exploitation. Treating properly this subject would require the definition of performance indexes that combine the dependability attributes and other quantities related to the quality of the LHC experiments. This subject goes beyond the aim of the thesis and it is not treated here.

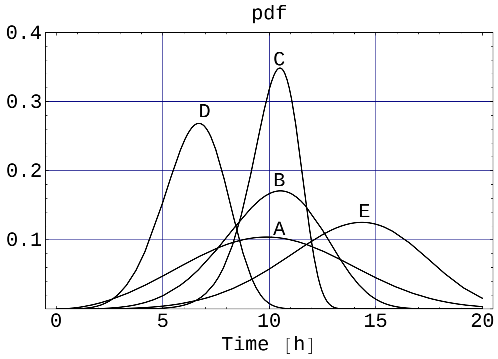


Figure 7.13: Weibull probability density functions (pdf) for the planned dump request with average 10 h and three different values of the variance (A,B,C) and for the planned dump requests with different averages (D,E).

Distribution	**Unsafety/y**	**FD/y**	M_{year}	T_M[hours]	σ
(A) Weibull [3, 11.3]	2.78×10^{-7}	4.083	399	10.018	3.57
(B) Weibull [5, 11]	2.58×10^{-7}	4.089	399	10.02	2.31
(C) Weibull [10, 10.61]	2.49×10^{-7}	4.087	399	10.028	1.25
(D) Weibull [5, 7]	1.51×10^{-7}	3.75	573	6.37	2.16
(E) Weibull [5, 15]	3.68×10^{-7}	4.27	306	13.67	9.95

Table 7.11: Sensitivity to the planned dump requests distribution.

7.5.3 Analysis of the Operational Scenario 3B

In this scenario, three independent sources of dump request are considered: the false dumps, the planned dumps and the beam induced dumps. The beam induced dump requests have a bigger likelihood during the injection, ramp and squeeze phase, roughly in the first two hours of the operational period when the beam is more unstable. After this period, they are expected to proportionally decrease to a constant rate, corresponding to the stable physics conditions in the LHC. The distribution for the beam induced dump request is derived from (3.1) using as rate $\lambda(t) = \lambda_\infty + (t^b + \frac{1}{\lambda_0})^{-1}$, with $b > 1$. The rate $\lambda(t)$ starts at λ_0 and tends asymptotically to λ_∞. The probability density function resulting from the beam induced dump request (b = 3, $\lambda_0 = 0.1$ and $\lambda_\infty = 0.001$) and the planned dump request (Weibull[5,11]) is shown in Figure 7.14. The larger probability of terminating the operation in the first 2-3 hours is due to the 'beam instability'. If the system passes this first critical stage, then the planned dump requests become the main cause of dump request with the maximum likelihood around 10 h. The system unsafety calculated by equation (7.17) is 2.964×10^{-7} for 592 missions per year. The result is adjusted following steps 2, 3 and 4. The average mission time is 8.1 hours that implies 475 missions alternate to 2 hours checks for one year of operation. The recalculated unsafety results in 2.401×10^{-7} per year. Using formula (7.19), the 475 dump requests are apportioned to 361 planned dumps, 110.1 beam-induced dumps (23% of the total) and 3.9 false dumps.

Sensitivity to the Beam Induced Dump Requests

The sensitivity to the beam induced dumps distribution is analysed with respect to the three parameters b, λ_0 and λ_∞. If b = 2, the rate decreases slower to the asymptotic value. The unsafety calculated by (7.17)is 2.89×10^{-7} per year for 631 dump requests. If the initial value λ_0 is 1 instead of 0.1 (i.e. more beam instability), the unsafety is 2.72×10^{-7} with 1380 dump requests per year. If λ_∞ is 0.01 the unsafety becomes 2.93×10^{-7} with 620

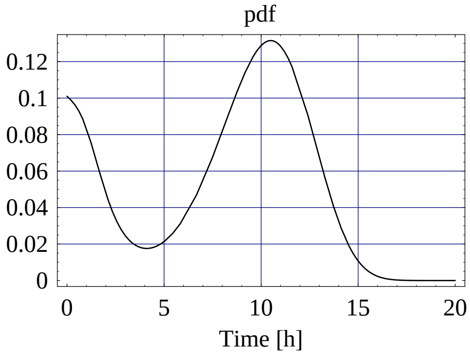


Figure 7.14: Probability density function for the external dump requests of the operational scenario OP3B.

	Unsafety/y	**FD/y**	**BI/y**	M_{year}	T_M [hours]
Default	2.401×10^{-7}	3.9	110.1	475	8.1
b = 2	2.295×10^{-7}	3.88	166.8	502	7.6
λ_0=1	1.725×10^{-7}	3.1	615	876	3.48
λ_∞=0.01	2.330×10^{-7}	3.9	146.7	492	7.74

Table 7.12: Sensitivity to the beam induced dumps distribution.

missions on average per year. Results are adjusted following step 3 and 4 and shown in Table 7.12 for unsafety, number of false dumps (FD), number of beam induced dumps (BI), number of missions (M_{year}) and average mission time (T_M). For the three cases unsafety is smaller than the unsafety obtained for the default OP3B. This can be explained by the fact that the mission average length is shorter than before, which is a consequence of the higher probability of terminating the operation with a beam induced dump, either in the early or in the late stage of the mission. The number of false dumps remains practically the same in all analysed cases.

7.6 Comparison Between the Different Operational Scenarios

The results of the analysed operational scenarios are shown in Table 7.13. The operational scenario OP3B is the most realistic and complete one. It results in a system unsafety of 2.401×10^{-7} and about 4 false dumps on average per year. These figures are also confirmed by the other scenarios. In particular, the value for unsafety calculated for OP3B is very close to that obtained for OP1, but for a shorter mission time. It is worthwhile reminding that OP1 enters the check phase only after 10 hours, even if the system has dumped the beam before with a false beam dump. On the contrary, OP3B enters the check phase as soon as a dump request has been generated. The two scenarios can be compared if the mission length of OP1 is set to 8.1 h, equal to the average mission length of OP3B. In this case, the resulting unsafety of OP1(2) is 1.88×10^{-7} per year, which is smaller than the value obtained for OP3B. The same holds for OP1 and OP3A, which have identical average mission length but different unsafety. Three factors concur in this result: 1) the fail unsafe rates that increase in time, 2) the number of missions per year, determined by the average mission length, and 3) the variance of the distribution of the mission length. For short average mission length the system is exploited less due to the two hours of check, and also operates at lower failure rates, with the result that the system is safer as demonstrated in 7.5.2. In the case considered, the difference between OP3B and OP1(2) or OP3A and OP1 cannot be explained by the average mission lengths, which are identical, but it is explained by the variance of the distribution of the mission length. Figure 7.15 (right) shows the distribution of the mission length for OP3A together with the fail unsafe rate ($\times 10^9$). For OP3A there exist a probability that the system works at a higher failure rate for $t > 10$, while this is not possible for OP1 that stays within 10 hours. The additional hazard depends on the variance of the distribution of the mission length. If the variance tends to zero, the two scenarios are practically identical, and

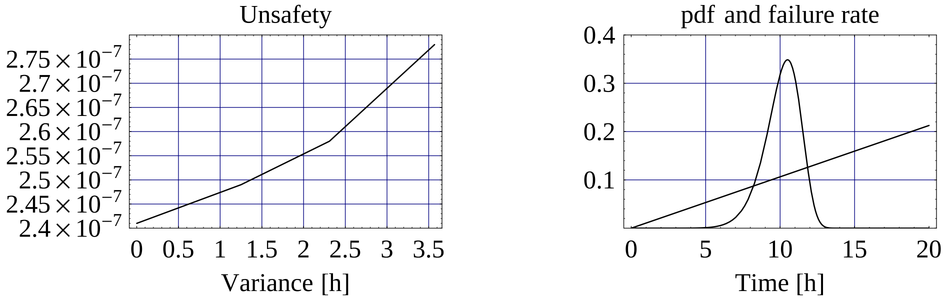


Figure 7.15: The failure rate ($\times 10^9$) and the distribution of the planned dump requests (right). The unsafety with respect to the variance of the planned dump requests (left).

results should coincide. The validity of this thesis is attested by plot of Figure 7.15 (left). For a variance tending to zero, also unsafety decreases and tends asymptotically to the value obtained for OP1, which corresponds to a zero variance distribution (i.e. a delta of Dirac in t = 10h).

The obtained figures for unsafety and number of false dumps for all presented scenarios can be doubled[13] for the two LBDS in the LHC. The analysis could be extended over the many years until machine disposal but this would require a deeper understanding of the wearing and aging processes, which have not been treated in the present study where all component failure rates are assumed constant. Assuming an ideal overhaul maintenance between the operational years, with perfect replacement of all aged parts, the system unsafety after N years would be N times the unsafety calculated for one year of LHC operation.

7.7 Reliability Runs

Reliability runs of the LBDS are planned before the commissioning of the LHC with beams. The aim of the reliability runs is twofold: 1) to

[13] This is possible because the LBDS are assumed to be identical and their failure processes are assumed to be statistically independent.

Scenario	**Missions/y**	T_M[hours]	**Unsafety/y**	**False dumps**
OP1	400 (fixed)	10.0 (fixed)	2.418×10^{-7}	4.0
OP1(2)	475 (fixed)	8.1(fixed)	1.880×10^{-7}	3.9
OP2	400 (fixed)	10.0 (fixed)	3.150×10^{-5}	4.0
OP3A	399 (average)	10.02(average)	2.580×10^{-7}	4.1
OP3B	475 (average)	8.1 (average)	2.401×10^{-7}	3.9

Table 7.13: Summary of the analysis of the operational scenarios.

troubleshoot the infant mortality problems and 2) to validate the reliability figures obtained by the presented analysis of the system. As a sample study, a reliability testing policy for the MKD generators is presented for the point 2, and formulated like a hypothesis to be verified (i.e. accepted or rejected):

- (H1) The MKD generator branch failure rate is $\lambda \leq 10^{-4}$/h.

The hypothesis sets an upper limit of the failure rate equal to 10^{-4}/h, which corresponds to a safety level of SIL3 for the system of 15 MKD assemblies.

The mathematics of testing relies on statistical inference [39, 80, 87]. The lifetime of a component is estimated from the number of observed failure within a population of N_C identical components tested for a time T. In the analysed example, the population is the set of 60 independent identical MKD pulse generators branches. A single test experiment consists of arming the 60 branches and then pulsing according to the sequence that will be repeated many time during the LHC operational cycle. The following assumptions define the test policy:

- (i) The test interval T consists of experiments of fixed duration τ during which the branches are armed waiting in stand-by for the pulse.

- (ii) The effective time for test is assumed to be 75% of the total. The remaining 25% is downtime.

- (iii) The generator branch failure rate λ is assumed to be constant and identical either while pulsing or in stand-by.

- (iv) Failures are discovered by inspection, after pulsing, and the failed components are replaced to keep a 60 units population.

At the end of the test period, the number k of observed failures over $N_E = 60T/\tau$ experiments is a Binomially distributed random variable:

$$P(k,T) = \binom{N_E}{k} p^k (1-p)^{N_E-k} \tag{7.20}$$

where $p = 1 - R(\tau) = 1 - e^{\lambda\tau}$ and $0 \leq k \leq N_E$. The average failure rate is calculated by the maximum likelihood estimator:

$$\hat{\lambda} = \frac{-\ln(1-\frac{k}{N_E})}{\tau} \tag{7.21}$$

The goal of the test is to accept/reject the hypothesis H1, with a certain confidence level on the decision threshold[14]. The one-sided confidence interval of the generator branch failure rate λ_{H1} is a function of the number of observed failures k, the given confidence level c and a test duration T. It is calculated by solving the integral equation below:

$$\int_0^{\lambda_{H1}} P(k,T)d\lambda = c \int_0^{\infty} P(k,T)d\lambda \tag{7.22}$$

As an alternative, the solution can be obtained using the χ^2_n distribution [80] that is a special case of the gamma distribution:

$$\chi^2_n = \gamma_{[1/2,n/2]} = \frac{t^{(n/2-1)}e^{-t/2}}{(2^{n/2}\Gamma_{n/2})} \tag{7.23}$$

where $\Gamma_{n/2} = \int_0^{\infty} x^{n/2-1}e^{-x}dx$ and n is the degree of freedom of the distribution. The (7.23) approximates the binomial distribution provided that $p < 0.1$, which is verified for the case studied. The following inequality holds:

$$p < \frac{1}{2N_E}\chi^2_{2(k+1)c} \tag{7.24}$$

from which it is possible to deduce λ_{H1}.

[14] In literature, this test is called type I censored life testing [39].

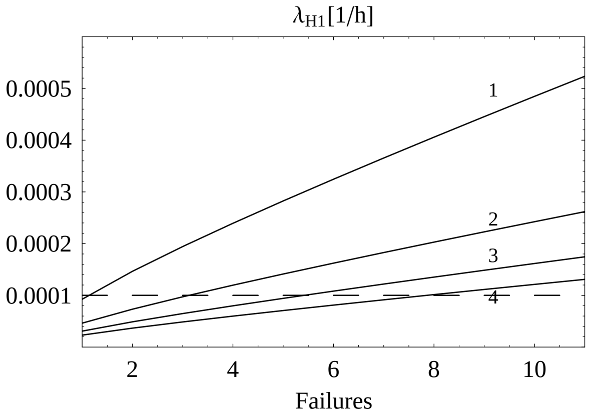


Figure 7.16: The 95% one side confidence interval for the MKD generator branch failure rate.

The one-sided confidence interval curves versus the number of observed failures $k = [0 \ldots 10]$ are shown in Figure 7.16 for $c = 0.95$, the length of a single experiment $\tau = 1$ h and the test time $T = 1$, 2, 3 and 4 months. The curves cross the decision threshold (10^{-4}/h) for a certain number of observed generator branch failures, above which the hypothesis H1 is rejected. The crossing point is one failure for one month of testing, while it is 6 failures for 3 months of testing. The latter is a large enough decision threshold and for this reason is chosen as test period for the MKD reliability run.

The time distribution of failures has been ignored so far, whereas it might be important, especially for assessing a reliability growth in the components under test. If failures are concentrated in the early testing period and once discovered they progressively disappear, then the system is said to experience a reliability growth. In this case, even if the acceptance threshold is globally exceeded, the hypothesis H1 might still be accepted. The choice of using a test period of three months at least could also cover these effects.

A second hypothesis could be formulated for verifying the existence of common mode failures in the MKD branches. This failure mode has not been included in the analysed models. The 60 branches may be tested separately and then the results compared to the results of the test on the 30 generators, each working with two branches in parallel. Eventual disagreement in the estimated statistics would attest a positive dependence of the components failure modes, and the consequent rejection of the assumed hypothesis of independence. A model for common mode failure has been described in section 5.1. The two branches are assumed to fail independently with rate λ but, as one of the two fails, the survived component will suffer from an additional rate, passing to the new failure rate $\lambda + \lambda_C > \lambda$. This effect can be observed using the same test policy.

The presented reliability test can be applied to the other systems of the LBDS provided that the decision thresholds are adjusted to the component under test.

Chapter 8

Dependability Assessment of the Machine Protection System

The results of the dependability analysis of the LBDS described in Chapter 7 are integrated into a simplified model of the LHC Machine Protection System, including the most important protection systems.

8.1 A Simplified Machine Protection System

The LHC Machine protection System (MPS) has been presented in Chapter 1. Here a simplified MPS is considered including the most important components that are involved in the protection task [33] that are: the Beam Interlocking System BIS, the LHC Beam Dumping System, the Beam Loss Monitors System (BLMS), the Quench Protection System (QPS) and the

System	**Quantity**	**Position in the LHC ring**	**Function**
LBDS	2	Sector 6	Beam extraction
BIC	16	2 per sector	Beam permit transmission
PIC	36	several per sector	Power permit
BLM	3500	several per sector	Beam loss detection
QPS	4000	several per sector	Quench protection

Table 8.1: The components of the simplified MPS.

Powering Interlock Controllers (PIC), see Table 8.1. In the simplified model, the BIS consists of 16 Beam Interlock Controllers (BIC) for the generation of the beam permit signal and the reception of the beam dump requests, see also section 1.3. The LBDS consists of two extraction systems, one per LHC beam. The BLM system consists of 3500 monitors (i.e. ionization chambers with preset detection threshold) plus electronics (i.e. VME crates) for the generation of the user permit signal [27]. The QPS consists of some 4000 channels (i.e. quench detectors and electronics) that survey any resistive transition of the superconducting magnets and can be activated to dissipate safely their stored energy [90]. The PICs, 36 in total, implement the power permit loop by interlocking the status of the power converters of the superconducting magnets in a similar way to the beam permit loop [75]. The power permit loop is established at the start of the operation and may be cut indirectly by the QPS via the PIC, resulting in a dump request to the local BIC. No difference in the configuration of the MPS for the 8 sectors of the LHC is assumed.

These systems are expected to cover all main types of hazard in the LHC, because a beam loss is the ultimate consequence of any critical failure in the machine. They also partially overlap their protection actions. For instance, a superconducting magnet quench may be caused by a slow beam loss and the same beam loss may be detected by different BLMs in the ring, thus increasing the overall coverage. These important features are included in the present study.

8.2 A Dependability Model for the Simplified MPS

Safety and availability are the dependability attributes of interest for the MPS and are represented in the state transition diagram of Figure 8.1, for a single LHC operation. The model consists of four states: the ready state (waiting for a dump request), the available state (after successful operation),

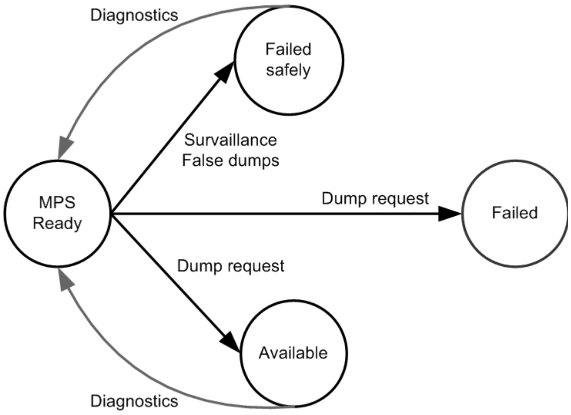


Figure 8.1: The simplified state transition diagram of the MPS.

the failed safe and failed unsafe state. The failed safe state is reached due to a detected failure in one of the MPS components, resulting in a false beam dump. The system is recovered to the ready state, after post mortem diagnostics, from all states with exception of the failed unsafe state that is absorbing.

The safety over one LHC operation (i.e. mission) is the probability that the system has dumped safely, whatever the cause of the dump request. The availability is the probability the system has dumped safely at a planned dump request or at a detected beam loss in the LHC, and excludes the false dump requests generated internally to the MPS.

Two different combinatorial models are chosen for modeling separately the safety and the availability of the simplified MPS. The model for **safety** is shown in Figure 8.2, in which the simplified MPS is arranged into a reliability (safety) block diagram. The blocks represent the components of the MPS that must be available at the time of the dump request, in the order in which they are involved, from the detection of the hazard, to the distribution of the dump request to the LBDS for the beam extraction. With the exception of the BIC and the LBDS, which are always required, the other MPS systems may be demanded or not depending on the source of the dump request. Four different sources of dump request are considered in the model:

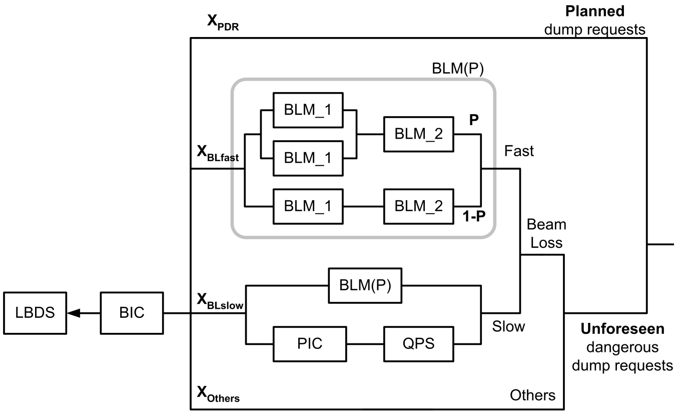


Figure 8.2: The reliability block diagram of the simplified MPS.

- X_{PDR}: planned dump requests from the control room
- X_{BLfast}: fast beam losses (< 10 ms)
- X_{BLslow}: slow beam losses (> 10 ms)
- X_{Others}: other sources.

False dumps have been assumed safe and for this reason they do not enter the list. The model also accounts for the possibility that more systems work in cross-redundancy contributing in parallel to the generation of the same dump request. This holds for the BLM and the QPS (via the PIC) that work in cross redundancy for slow beam losses, which also results in a magnet quench. Redundancy is also considered within the BLM. The BLM detectors with the front-end electronics (BLM1) are separated from the VME crate electronics (BLM2) that collect their signals and transmit the dump request. A constant P ($0 \leq P \leq 1$) accounts for the probability that two monitors detect the same beam loss. Redundancy within more BLMs is possible but not taken into consideration.

The formula for the calculation of safety of the simplified MPS is:

$$\begin{aligned} S = \; & S_{LBDS}S_{BIC}[X_{PDR} + X_{others} + X_{BLfast}S_{BLM}(P) \\ & + X_{BLslow}(1 - (1 - S_{BLM}(P))(1 - S_{PIC}S_{QPS}))] \end{aligned} \tag{8.1}$$

$$S_{BLM}(P) = PS_{BLM2}(2S_{BLM1} - S_{BLM1}^2) + (1 - P)S_{BLM1}S_{BLM2}$$

where S_x stands for the safety of the system x, and $X_{PDR} + X_{BLfast} + X_{BLslow} + X_{others} = 1$ are the relative contributions from the four sources of dump requests. In (8.1) each branch contributes to the total unsafety by the weight of the respective fraction of dump requests. For N consecutive missions the safety of the MPS is S^N.

The model for the **availability** of the simplified MPS is the series of all its components. The false dumps of just one of them is sufficient to abort the operation. The overall MPS false dump rate λ_{FD} is the sum of all false dump rates:

$$\lambda_{FD} = 2\lambda_{LBDS} + 16\lambda_{BIC} + 36\lambda_{PIC} + 3500\lambda_{BLM} + 4000\lambda_{QPS} \tag{8.2}$$

The number of false beam dumps over N operations per year is a binomially distributed random variable. The formula (7.12) can be applied with $p = 1 - \exp(-\lambda_{FD}T)$, where T is the length of one LHC operation. The average number of false dump and the standard deviation are respectively $N \times p$ and $\sqrt{[N(1-p)p]}$. The number of false dumps generated in the MPS does not depend on the given apportionment of dump request and the cross-redundancy.

8.3 Analysis of the Simplified MPS

The following assumptions define the scenario I for the dependability analysis of the simplified MPS model:

- (A1) 400 missions, 10 hours each, alternate with 2 hours of check for 200 days of operation.
- (A2) The diagnostics during the check period has a different effect on the MPS components. For the LBDS this is assumed to be a regeneration point, see Chapter 7, while this is partially true for the BIC and the BLMs. The QPS and the PICs are inspected at periodic interval of one month or after a power abort.

System	Unsafety/year	False dumps/year	
		Average	Std. dev.
LBDS	$2.4 \times 10^{-7} \times 2 = 4.8 \times 10^{-7}$	$4 \times 2 = 8.0$	2.0
BIC [13]	1.4×10^{-8}	0.5	0.5
BLM [37]	1.44×10^{-3}(BLM1) 0.06×10^{-3}(BLM2)	17.0	4.0
PIC [97]	0.5×10^{-3}	1.5	1.2
QPS [90]	0.4×10^{-3}	15.8	3.9
MPS	2.3×10^{-4}	41.0	6.0

Table 8.2: Safety and number of false dumps for the MPS for the scenario I.

- (A3) Dump requests are apportioned in 60% planned beam aborts, 15% fast beam losses, 15% slow beam losses and 10% other sources[1].

- (A4) Cross-redundancy within the BLM is not included, that is $P = 0$.

For the BIC, the BLMS, the QPS and the PIC, the dependability figures have been obtained from different studies [13, 37, 90, 97]. The results for the individual systems and the complete MPS are shown in Table 8.2. Unsafety is 2.3×10^{-4} per year, which corresponds to an equivalent failure rate of 0.58×10^{-7} per hour which is SIL3.

A different scenario II assumes a larger contribution from the fast beam losses. The dump requests apportionment for this scenario is 20% planned, 45% fast beam losses, 25% slow beam losses and 10% others. The unsafety becomes 6.8×10^{-4} per year, which corresponds to an equivalent failure rate of 1.7×10^{-7} per hour that is SIL2, see Figure 8.3 (left). The result for safety can be explained by the fact that the fast beam losses are covered by only one BLM as specified in the assumption A4 ($P = 0$). This assumption is relaxed in another scenario III. The unsafety decreases with the parameter P to a minimum value of 2.8×10^{-5} per year for $P = 1$, see Figure 8.3 (right). This corresponds to an equivalent failure rate of 7.0×10^{-9} per hour that is SIL4.

[1] This distribution is in part inspired by the HERA accelerator experience [93].

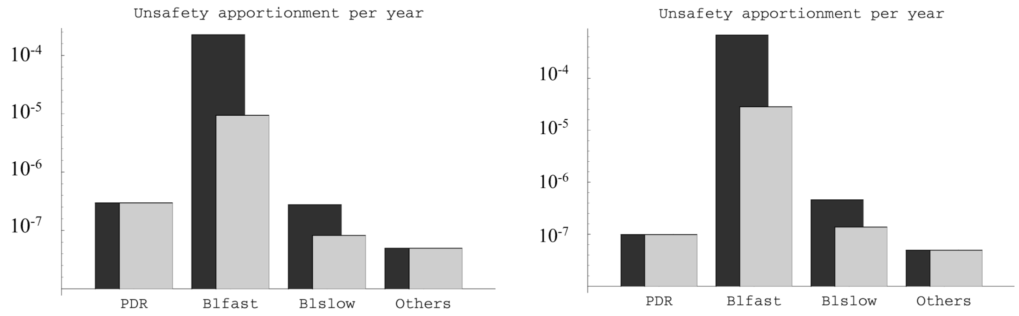


Figure 8.3: Sensitivity to dump requests apportionment (left): gray bars for scenario I, black bars for scenario II. Sensitivity to cross-redundancy within the BLM (right): black bars for scenario II, gray bars for scenario III with $P = 1$.

The number of expected false dumps is 41 (±6) per year for all scenarios, about 10% of the total machine fills. The average number of false dumps is not exactly the sum of the contribution of each system taken separately but a value slightly smaller as these act as concurrent events.

The three different scenarios I, II and III show a strong dependency of safety with respect to the source of beam dump request and the redundancy among the protection systems. Depending on the setting of the parameters in the model, the calculated safety of the simplified MPS ranges between SIL2 and SIL4. The number of expected false dumps is 41 false dumps per year, 10% of the total machine fills. The power supplies of the electronics for the different systems are expected to be the main cause of the false dumps, with a failure rate of $\cong 10^{-5}$/h taken from literature. An alternative design with two power supply units in parallel has been demonstrated effective for the QPS where the false dumps can be halved from 16 to 8 per year [90].

The presented model is suited to improvements. Only the LBDS, the QPS and the BLM have been analysed in details, while the results for the PIC and the BIC refer to simplified architectures and should be taken as provisional. Other protection systems like the Beam Current Transformer, the Beam Position Monitors, the collimators are not included in the model, but could contribute significantly to safety and the number of false dumps. An improved study could also take into consideration the integrity of the post mortem diagnostics and the rearming procedures that will play an important role not only for the safety of the LBDS, as demonstrated in this thesis, but also for the other systems.

Chapter 9

Conclusions

The most important results of the dependability study of the LHC Beam Dumping System are presented in this chapter together with some concluding remarks.

The calculated **safety** of the LBDS largely complies with SIL4 under the assumption that checks are able to regenerate the system to an ‘as good as new’ state, after every beam dump, while it can decrease to SIL3 if this assumption is removed. The largest contribution (99%) to unsafety is expected to come from the magnet assemblies MKD, MSD and MKB, with a remaining small fraction from the electronics. Looking per system the MKD is the most critical component of the LBDS, accounting for more than 75% of the total unsafety, followed by the MSD (18%) and the MKB (6%). Within the MKD, the increase of the magnet failure rate with time (e.g. due to wearing) has been shown to be a possible concern for safety. The system has also been demonstrated to be sensitive to redundancy and on-line surveillance, and without all or just part of these features the required safety would not be obtained.

The LBDS **unavailability** has resulted in 4 ± 2 false dumps per LBDS, per year. Again, the MKD system is responsible for the largest contribution with around 60% of the total generated number of false dumps, followed by the MKB (20%) and the MSD (11%). A significant fraction is expected to come from the power triggers of the MKD and MKB systems and results in 2 false dumps per year, 50% of the total.

The presented figures of safety and number of false dumps refers to an operational scenario that accounts for missions of random duration driven by three independent events: a false beam dump generated by a detected failure in the LBDS, a planned dump issued by the control room, and the beam induced dump requests generated by detected beam anomalies during the LHC operation. Other operational scenarios have been studied, which do not differ significantly in the obtained results.

Safety and availability are a trade-off for the LBDS. In most analysed cases they are well balanced, as the redundancy and surveillance are either strictly necessary or assure some additional safety margin that results only in a small contribution to the number of false dumps. Nevertheless, for some cases the failure of the components in the redundant architecture and in the surveillance can have an important impact on the system availability. For example, an increase of the failure rate of the two independent power trigger modules in the power trigger system has resulted a minor safety concern because of surveillance and redundancy, but a major concern for the availability as this significantly increases the number of false dumps. The same holds for the failure rate of the power supplies in the electronics and the data acquisition channels of the BETS. These examples could be taken into account for the final installation of the equipment as they might represent a potential weakness in the design trade-off.

The results of the study are going to be validated through reliability runs. A test period of three months is planned to collect the necessary statistics on failures of the MKD branches and to draw a decision about the reliability. Six generator branch failures observed in that period are the calculated decision threshold for accepting/rejecting the hypothesis that the MKD system is at least SIL3 with a 95% confidence level.

The dependability study of the LBDS has not only a relevance for the LHC but, due to the complexity of its architecture, it represents a benchmark for the applied methodology of modeling and analysis the dependability of complex systems. Some numbers may give an idea of the complexity of the analysed case. The studied LHC beam dumping system consists of the series of 15 MKD extraction kicker magnets, 15 MSD extraction septum magnets and 10 MKB dilution kicker magnets, with the associated power converters, the electronics for triggering and re-triggering the MKD and MKB systems, the beam energy measurement and the beam energy tracking systems. In total, about 2130 independent failure modes at component level have been identified for the LBDS architecture and arranged in 21 independent failure modes at system level. These failure modes have entered a compact state transition diagram of six states for the description of the system failure processes and the analysis the dependability attributes. A distinction is made between the collection of failure statistics at component level and the modeling of the failure processes at system level. For the first issue, the presented study applied Failure Modes Effects and Criticality Analysis (FMECA) together with international standards and military handbooks, with the aim of disposing of a homogeneous data source for the component failure modes and rates. For the description of the system failure processes, a modeling hierarchy is presented for supporting the analysis. This represents an innovative part of this thesis. Combinatorial techniques are used at lower level for deducing the statistics of failure modes while the mathematics of Markov processes and Markov regenerative processes is applied at higher level for representing the failure modes into a state transition diagram. The adopted approach is alternative to other more traditional modeling techniques (e.g. fault trees). The main advantages of this approach are listed below:

- The failure process is described at system level, with the component failure modes and statistics driving the state transitions.
- The phase transitions (i.e. missions and checks) can be embedded in the model and treated with the Markov regenerative mathematics. This permits to model separately the different events that can issue a dump request, giving the possibility of larger range of scenarios to be

analysed.

- The model is able to represent both safety and availability of the LBDS, which is another advantage of state transition diagrams with respect to a combinatorial approach, where only one dependability attribute can be addressed at a time.
- A compact state transition diagram is suited for presenting results in contexts (e.g. project management) where a global picture of the system is often preferred.

A simplified model of the Machine Protection System (MPS) including the LBDS, the Beam Interlock Controller, the Beam Loss Monitors, the Quench Protection System and the Powering Interlock Controller has been studied. The model has been analysed for different operational scenarios, which resulted in a safety level between SIL2 and SIL4, and 41 ± 6 expected false dumps per year, 10% of the total number of machine fills. The resulting safety has been demonstrated to be sensitive to the type of dump requests and their coverage, in particular those caused by the fast beam losses that are clearly the most critical events concerning safety in the LHC. The results can be refined for some of the components of the MPS, and the model can be extended by including other protection systems. Even without other improvements, this study has the merit of being the first attempt to address the dependability analysis of the LHC MPS in homogenous way, by including the most important components involved in the machine protection task.

With respect to the SIL3 requirement demanded from the MPS, the LBDS has been demonstrated to be one of the safest sub-systems, largely within SIL4, with a calculated contribution of 8 false dumps (4×2 LBDS), which is considered to be an acceptable fraction of the total number of the expected false dumps generated by the MPS.

Appendix A

Reliability Prediction of the LBDS Components

This appendix gives an overview of the methods used to obtain failure rates statistics for the LBDS components. The information on the failure rates of the various components was only for a small part available from the manufacturer and CERN historical databases, while the largest part came from the military handbook [66]. In the reliability handbook, these figures are given for accelerated-test conditions (T = 55 $°C$, and 90% confidence level). They have been adjusted for a ground-fixed environment, using the formula $\lambda_a = \lambda_b \pi_Q$ where λ_b is the base failure rate π_Q is the quality factor. After that, failure rates have been apportioned into failure mode rates using manuals and standards [67] and [34]. Failure modes have been deduced at component level using the electrical layout. If not differently specified, they have all been assumed statistically independent.

A.1 A Sample Case Study: The MKD System

The reliability prediction is applied to calculate the failure rate statistics in the MKD systems, at component level. One MKD generator assembly is described in section 2.2, and shown in Figure 2.3 .

The primary switch layout is shown in Figure A.1. Two failure modes exist, SP1 silent (i.e. it fails to close at the input trigger) and SP2 erratic (i.e. it closes without an input trigger), see Table 6.2. A component may fail silent if at least one coaxial input (coax1, coax2) has failed short or both the diodes of the input section (D-IN) have failed open. The switch may fail erratically due to a failure in the GTO (Gate Turn-Off thyristor) stack that results in a voltage drop at the voltage dividers VDP1 and VDP2, detected by the retriggering system. In this sense, a breakdown in the resistor R1-10, the short of one GTO T1-10 and the leakage in the capacitors C1-10 do not result into a dump request. These failure modes are not detected at the voltage divider but at the resistor RN10. The failure is not critical because the survived GTOs can still drive the proper current and the dump request is not generated. The high-voltage stress at the secondary transformer (right branch of Trk in Figure A.1) is more serious and may provoke a short towards the primary one (left branch of Trk in Figure A.1) and then backwards to the power triggers. This failure is caught by the re-triggering system, not at the switches but at the primary capacitors as a voltage drop. The study of the back-propagation of the fault (sneak circuit analysis) and the consequences are still to be addressed and might suggest some extra-protection in the design. In total, the rate of the failure mode SP1 is 1.4 FIT (Failures In Time, 10^{-9}/h) and SP2 is 180 FIT, which is mainly due to the Tr transformer short. If a beam dump request were generated by a GTO failure then SP2 would be 2364 FIT, which turns to be 4464 FIT if the failure of one capacitor in the stack is included as well. This hypothesis is not taken into account.

The compensation switch layout is shown in Figure A.2. The failure modes are SC1 silent and SC2 erratic, listed in Table 6.2. The components may fail silent if the coax 3 has failed (either open or short) or the GTO T11 has failed open or one among R11, R12, and D2 has failed short. Other failures, like the opening of the diode stack RD1 (the combination of diode and resistor opening), are deemed very unlikely and therefore neglected. The switch may fail erratically for the erratic closing of the GTO T11. In total the rate of the failure mode SC1 is 309 FIT and SC2 is 216 FIT.

The primary capacitors assembly is shown in Figure A.3. Two failure modes exist, CP1 (capacitance leakage) and CP2 (connection failure), see Table 6.2. A capacitance drift can be caused by slow leakages that are detected by the BET and results in a dump request. The capacitor may fail open for the bad connection to the circuit. The opening of RDP1 is neglected. The rate of the CP1 failure mode is 270 FIT and CP2 is 30 FIT.

The capacitor assembly of the overshoot1 and overshoot2 circuits are shown together in Figure A.4. Two failure modes exist for the overshoot1: COS11 (capacitance leakage) and COS12 (connection failure), see Table 6.2. The OS1 capacitor may have a capacitance leakage that is detected by the BET and results in a dump request. The assembly fails open if COS1 has failed open or R-OS1 has failed open. The rate of COS1 failure mode is 270 FIT and COS2 is 61 FIT. Two failure modes exist for the overshoot2, COS21 (capacitance drift) and COS22 (open), see Table 6.2. It fails open if the capacitor COS2-A/B has failed open. The primary capacitor may have a capacitance leakage that is detected by the BET and results in a dump request. The open failure of both diodes DC-A and B can disconnect the OS2 capacitors from the compensation switch. This contribution is very unlikely, nevertheless it could accumulate undetected. The short of one (DC-A or B) of these diodes will produce a wrong powering and therefore a dump request. The rate of the COS22 failure mode is 30 FIT and the COS21 is 390 FIT (120 FIT from the DC-A, B).

The primary power supply circuit is shown in Figure A.5 (left). Two failure modes exist, PSP1 (under-voltage) and PSP2 (over-voltage), see Table 6.2. The PSP1 failure mode is due to the coax-IN failed open, one of the two diodes D-PSP failed (either short or open), the resistance R-PSP failed open or the capacitor breakdown. The result is a voltage drop detected by the BET and generating a dump request. In addition, the primary power supply can fail internally producing a voltage either lower or higher than expected.

This failure rate λ_{PS} is provided by the manufacturer (Heinzinger 35KV, 103 FIT). It is equally apportioned into 50% under voltage and 50% over voltage. In both cases the BET detects the error. The first is tolerated by redundancy while the second is not (over-voltage). The failure rate of PSP1 is $\lambda_{PS}/2 +$ 246.6 FIT and PSP2 is $\lambda_{PS}/2$.

The overshoot2 power supply circuit is shown in Figure A.5 (right). The component may fail in the input section due to the coax-IN failed open, the failure of D-OS2 (either short or open) and the short of R-OS2. In addition, the OS2 power supply can fail internally. Its failure rate λ_{OS2} is provided by the manufacturer (Heinzinger 300V, 103 FIT). These failures mode are caught by the BET. The failure rate of PSOS2 is λ_{OS2} + 66.4 FIT. The overshoot1 power supply circuit may fail due to the coax-N failure or the failure of the power supply itself (Heinzinger 350V, 103 FIT). This failure is not monitored and does not generate a dump request. The failure rate of PSOS1 is λ_{OS1} + 5.2 FIT.

The magnet is considered as the assembly of the coil plus transmission cables and connectors to the generator, see Figure A.6. It may fail (not powered) due to a short in one of the 8 coaxial cables (16 sockets) each of failure rate λ_{TX}. The open failure of capacitor Cm and the failure of Rm2 result in the propagation of the current back to the pulse generators with unpredictable consequences. At present, the rate λ_L of the coil failure (open/short) is unknown. The failure mode M, including the transmission lines, has rate $8\lambda_{TX} + \lambda_L$. with assumed $\lambda_{TX} = 10$ FIT and $\lambda_L = 100$ FIT.

The failures in the diode stacks of the primary capacitor, in the compensation switch assembly and the diodes DC-A,B may accumulate silently during the operation. The possibility of detecting these failures during post mortem, as a secondary effect on the pulse shape, is not trivial and it is not considered at moment. Failure rates and modes are shown in Table A.1. For each components a quality factor and a base failure rate are given, which are used to calculate the failure rate adjusted according to the parts count method of the MIL HDBK 217F, see also 3.2. The failure rate are apportioned into failure modes.

A.2 The LBDS Components Failure Rates

The reliability predictions have also been conducted for the other systems in the LBDS. The results are presented in Tables A.2 to A.10 for the MKD system, the MKB system, the MSD system, the power triggers, the triggering system, the re-triggering system, the BETS and the BEMS. Most failure modes are given a failure rate deduced from literature. All failure rates not deduced from literature have been left unspecified with their symbol. Then a value for the analysis has been assumed looking at similar components and/or technology. For example, the power supplies of the MKD and MKB have an assumed failure rate (λ_{PS} and λ_{OS1}) of 1000 FIT. The voltage dividers are given a rate (λ_{VD}) of 100 FIT, the magnet coil is given a failure rate (λ_L) of 100 FIT and the transmission lines a rate (λ_{TX}) of 10 FIT. For the synchronization surveillance unit of the triggering system, as no information was available on literature about the apportionment of the failure, this has equally been apportioned to the safe and the silent mode.

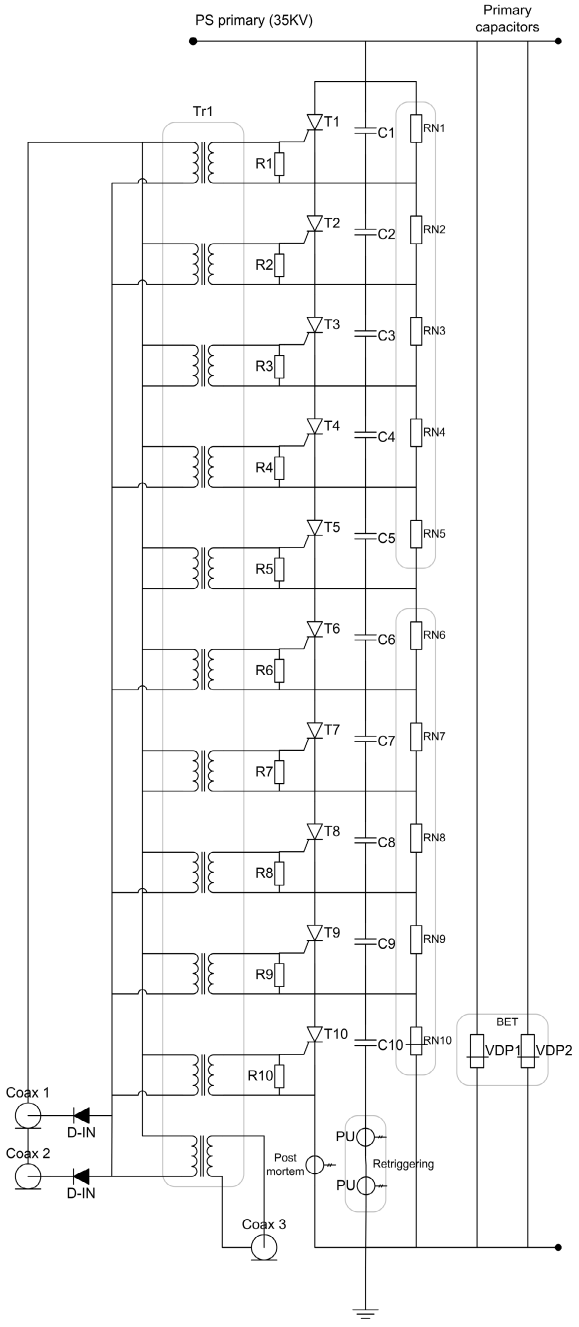


Figure A.1: The primary capacitor of the MKD generator

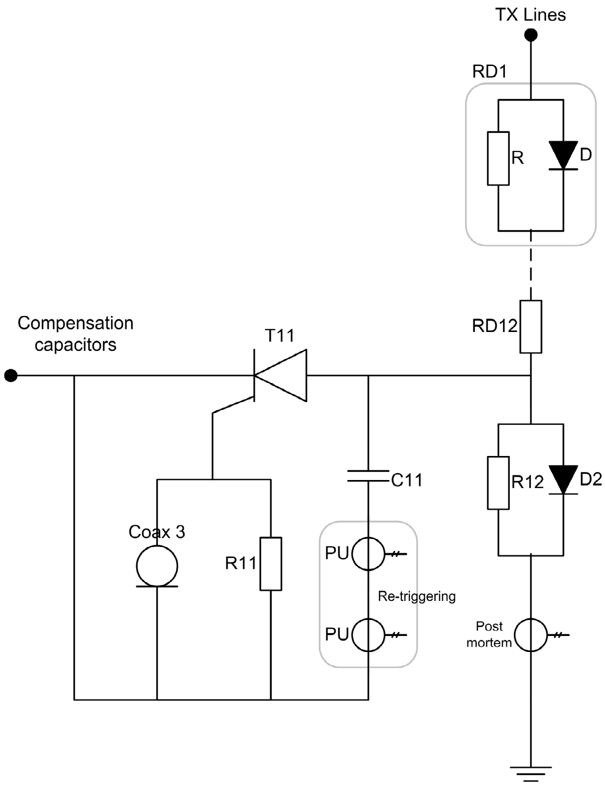


Figure A.2: The compensation switch.

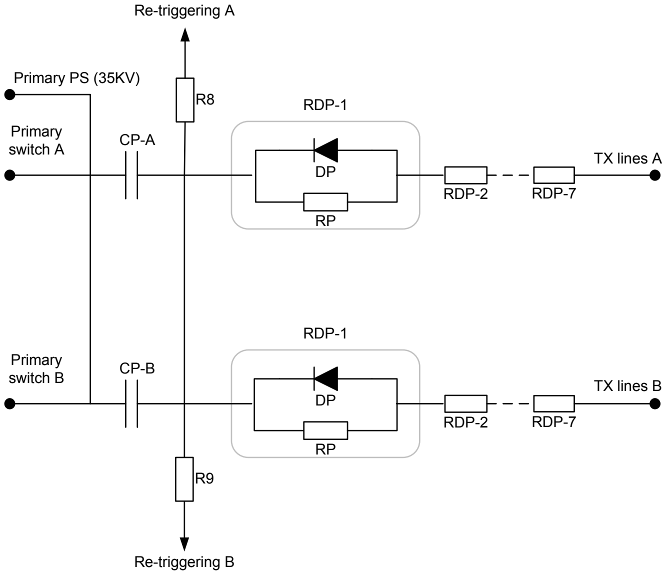


Figure A.3: Primary capacitor assembly.

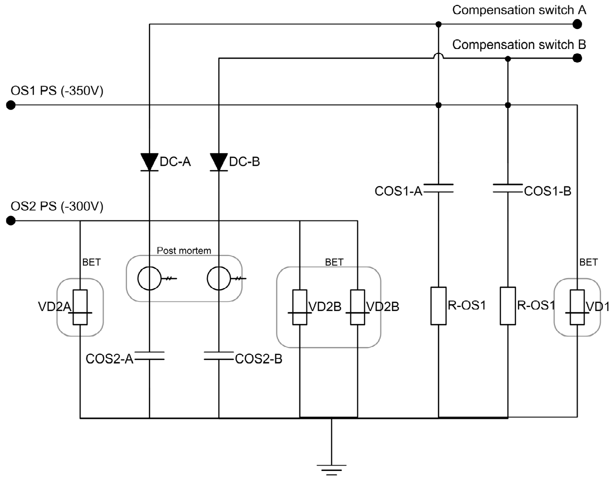


Figure A.4: Compensation capacitors assembly of the overshoot1 circuit.

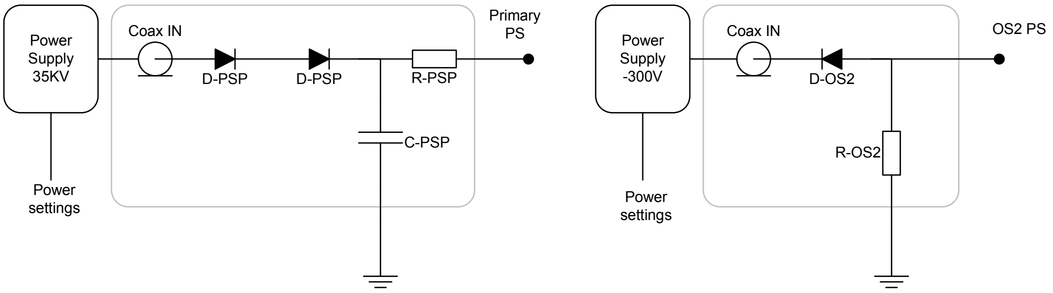
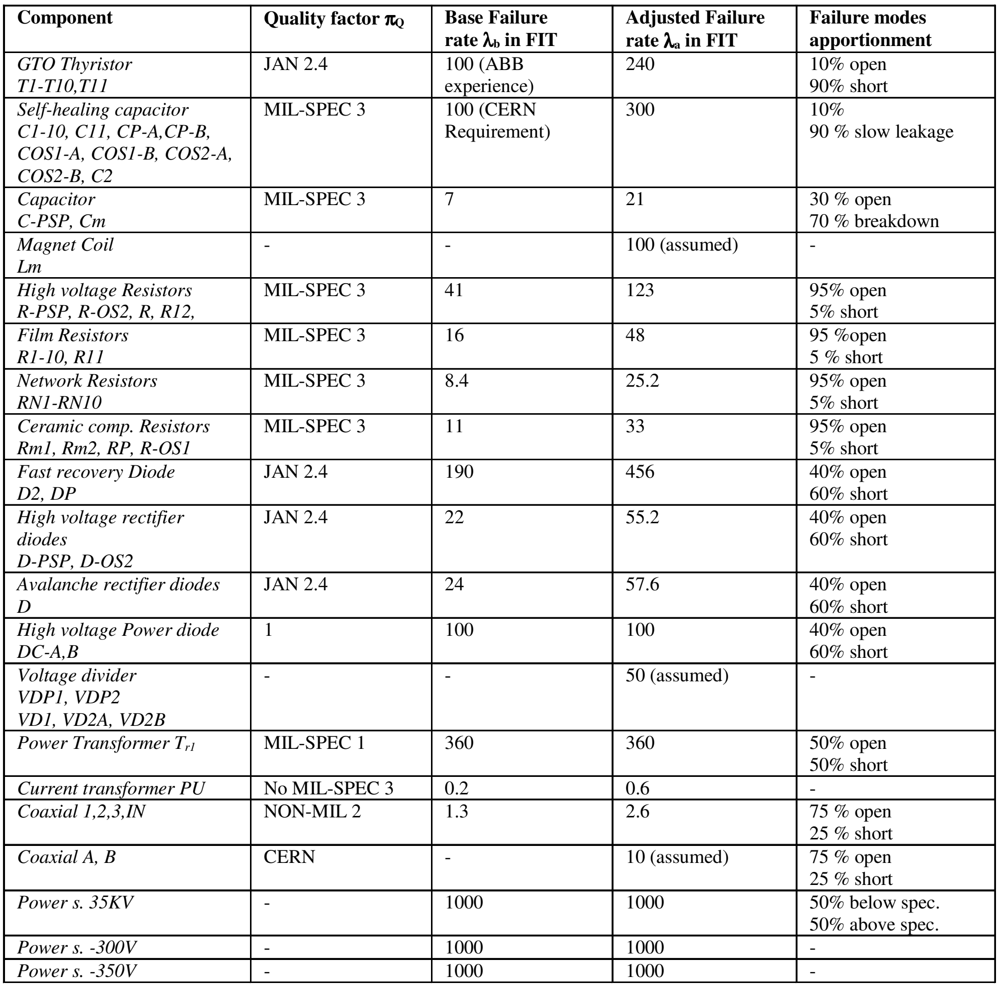


Figure A.5: Primary power supply circuit (left) and overhoot2 power supply circuit (right).



Component	**Quality factor π_Q**	**Base Failure rate λ_b in FIT**	**Adjusted Failure rate λ_a in FIT**	**Failure modes apportionment**
GTO Thyristor *T1-T10,T11*	JAN 2.4	100 (ABB experience)	240	10% open 90% short
Self-healing capacitor *C1-10, C11, CP-A,CP-B, COS1-A, COS1-B, COS2-A, COS2-B, C2*	MIL-SPEC 3	100 (CERN Requirement)	300	10% 90 % slow leakage
Capacitor *C-PSP, Cm*	MIL-SPEC 3	7	21	30 % open 70 % breakdown
Magnet Coil *Lm*	-	-	100 (assumed)	-
High voltage Resistors *R-PSP, R-OS2, R, R12,*	MIL-SPEC 3	41	123	95% open 5% short
Film Resistors *R1-10, R11*	MIL-SPEC 3	16	48	95 %open 5 % short
Network Resistors *RN1-RN10*	MIL-SPEC 3	8.4	25.2	95% open 5% short
Ceramic comp. Resistors *Rm1, Rm2, RP, R-OS1*	MIL-SPEC 3	11	33	95% open 5% short
Fast recovery Diode *D2, DP*	JAN 2.4	190	456	40% open 60% short
High voltage rectifier diodes *D-PSP, D-OS2*	JAN 2.4	22	55.2	40% open 60% short
Avalanche rectifier diodes *D*	JAN 2.4	24	57.6	40% open 60% short
High voltage Power diode *DC-A,B*	1	100	100	40% open 60% short
Voltage divider *VDP1, VDP2* *VD1, VD2A, VD2B*	-	-	50 (assumed)	-
Power Transformer T_{r1}	MIL-SPEC 1	360	360	50% open 50% short
Current transformer PU	No MIL-SPEC 3	0.2	0.6	-
Coaxial 1,2,3,IN	NON-MIL 2	1.3	2.6	75 % open 25 % short
Coaxial A, B	CERN	-	10 (assumed)	75 % open 25 % short
Power s. 35KV	-	1000	1000	50% below spec. 50% above spec.
Power s. -300V	-	1000	1000	-
Power s. -350V	-	1000	1000	-

Table A.1: Failure rates and the apportionment of failure modes for the MKD components.

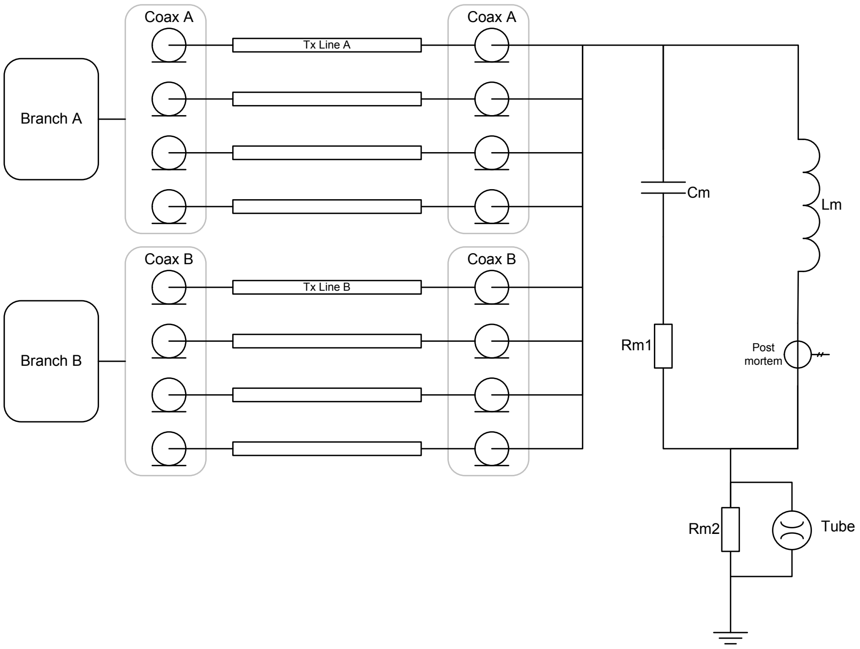


Figure A.6: Magnet coil and transmission lines.

Symbol	Failure mode	Failure rate/h FIT
PSP_1	Primary PS under-voltage.	$\lambda_{PS}/2 + 246.6$
PSP_2	Primary PS over-voltage	$\lambda_{PS}/2$
PSOS1	Overshoot1 PS failure	$\lambda_{OS1} + 5.2$
PSOS2	Overshoot2 PS failure	$\lambda_{OS2} + 66.4$
CP_1	Primary capacitor slow leakage	270
CP_2	Primary capacitor open failure	30
$COS1_1$	Overshoot1 capacitor slow leakage	270
$COS1_2$	Overshoot1 capacitor open failure	61
$COS2_1$	Overshoot2 capacitor slow leakage	391
$COS2_2$	Overshoot2 capacitor open failure	30
SP_1	Primary switch open	1.4
SP_2	Primary switch erratic	180
SC_1	Compensation switch open	309
SC_2	Compensation switch erratic trigger	216
M	Magnet failure	$10\lambda_{TX} + \lambda_L$
VD	Voltage divider failure	λ_{VD}

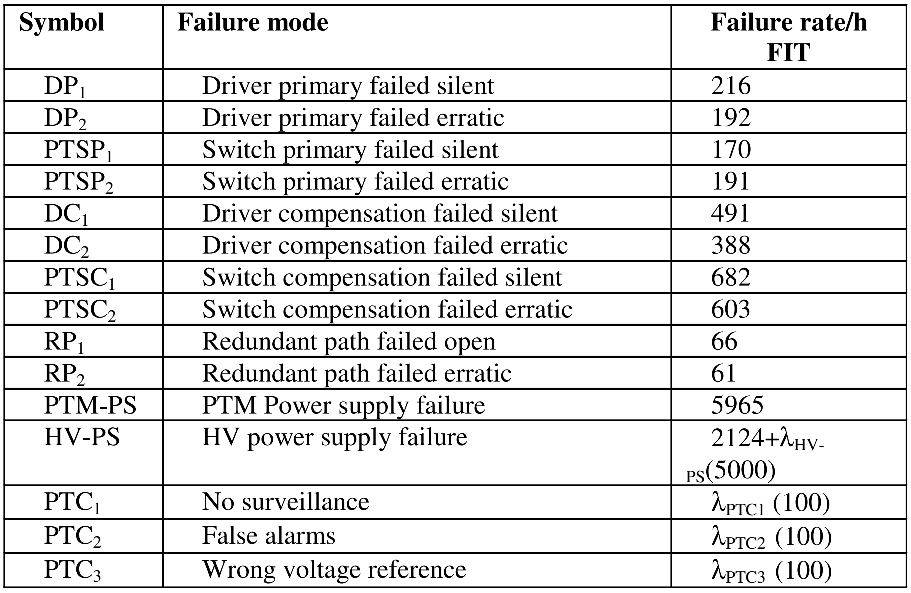
Table A.2: MKD system failure rates.

Symbol	Failure mode	Failure rate/h FIT
PS-H_1	Power supply failure (35kV)	$0.75 \times \lambda_{PS}$
PS-H_2	Power supply failure (35kV) – above 50%	$0.25 \times \lambda_{PS}$
PS-V_1	Power supply failure (–350kV)	$0.75 \times \lambda_{PS}$
PS-V_2	Power supply failure (–350kV) – above 50%	$0.25 \times \lambda_{PS}$
CH_1	Capacitor charging failure	$294 + \lambda_{VD}$
CH_2	Capacitor silent failure	30
CV_1	Resonant circuit charging failure	$294 + \lambda_{VD}$
$CV1_2$	Resonant circuit silent failure	$300 + \lambda_{VD}$
SW_1	Primary switch open	1.4
SW_2	Primary switch erratic	180
M	Magnet failure	$10\lambda_{TX} + \lambda_L$
VD	Voltage divider failed	λ_{VD}

Table A.3: MKB system failure rates.

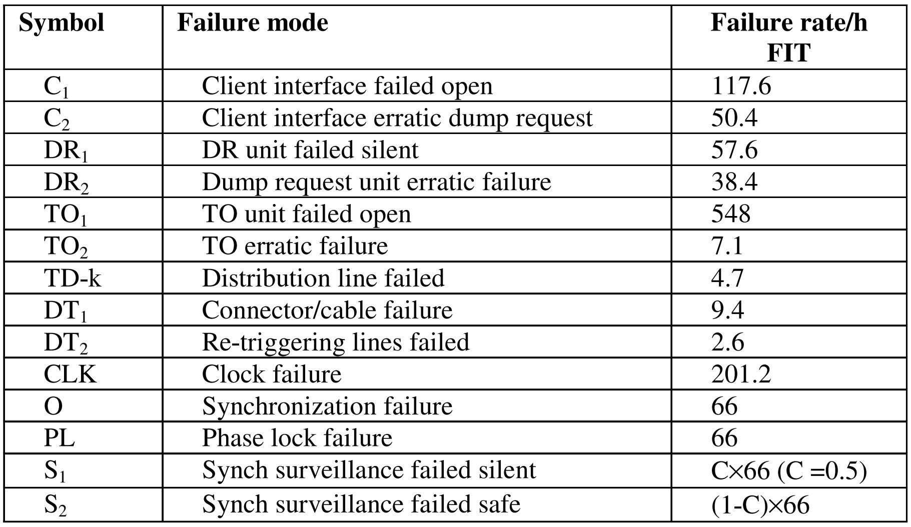
Symbol	Failure mode	Failure rate/h FIT
PC_1	Slow failures (> ms)	5000
PC_2	Fast failures	5000
M_1	Winding shorts or false contacts	100
M_2	Minor shorts (tolerated)	100
PLC	Unavailable	100
TS_1	Failed stuck-at	100
TS_2	Failed erratic	100
DCCT	Failed stuck-at	100
FC_1	Unavailable	100
FC_2	Failed safe, false alarm	100000
VD	Failed	λ_{VD}

Table A.4: MSD system failure rates.



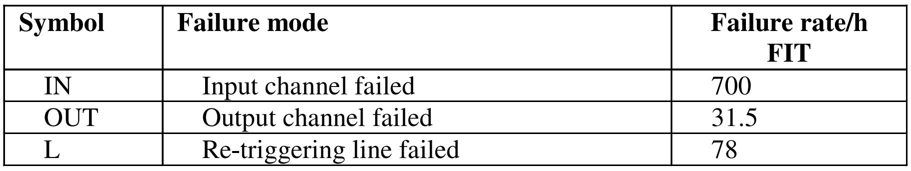
Symbol	Failure mode	Failure rate/h FIT
DP_1	Driver primary failed silent	216
DP_2	Driver primary failed erratic	192
$PTSP_1$	Switch primary failed silent	170
$PTSP_2$	Switch primary failed erratic	191
DC_1	Driver compensation failed silent	491
DC_2	Driver compensation failed erratic	388
$PTSC_1$	Switch compensation failed silent	682
$PTSC_2$	Switch compensation failed erratic	603
RP_1	Redundant path failed open	66
RP_2	Redundant path failed erratic	61
PTM-PS	PTM Power supply failure	5965
HV-PS	HV power supply failure	2124+λ_{HV-PS}(5000)
PTC_1	No surveillance	λ_{PTC1} (100)
PTC_2	False alarms	λ_{PTC2} (100)
PTC_3	Wrong voltage reference	λ_{PTC3} (100)

Table A.5: Power trigger failure rates.



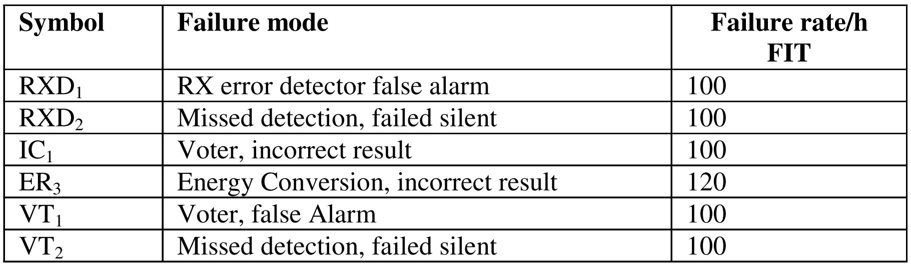
Symbol	Failure mode	Failure rate/h FIT
C_1	Client interface failed open	117.6
C_2	Client interface erratic dump request	50.4
DR_1	DR unit failed silent	57.6
DR_2	Dump request unit erratic failure	38.4
TO_1	TO unit failed open	548
TO_2	TO erratic failure	7.1
TD-k	Distribution line failed	4.7
DT_1	Connector/cable failure	9.4
DT_2	Re-triggering lines failed	2.6
CLK	Clock failure	201.2
O	Synchronization failure	66
PL	Phase lock failure	66
S_1	Synch surveillance failed silent	C×66 (C =0.5)
S_2	Synch surveillance failed safe	(1-C)×66

Table A.6: Triggering system failure rates.



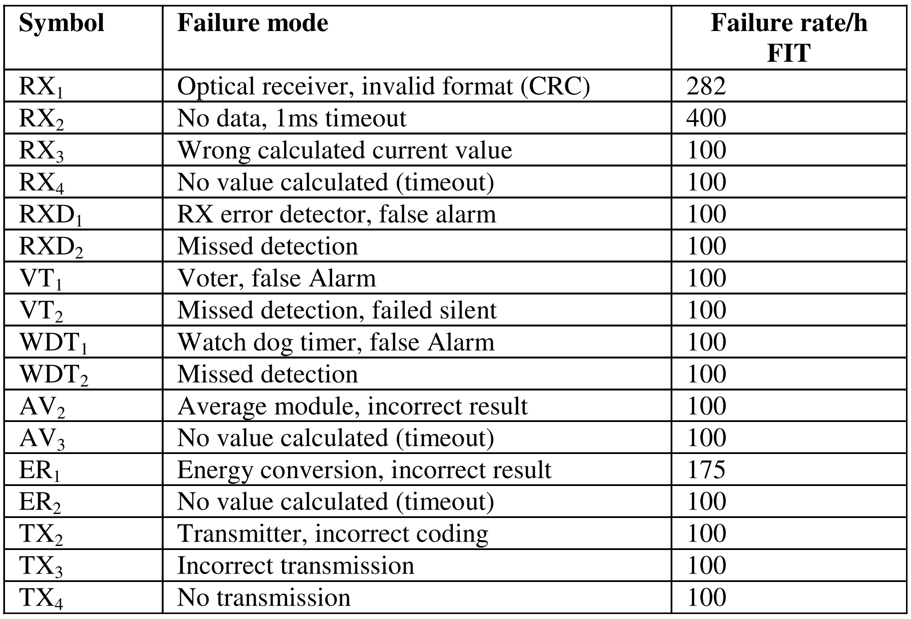
Symbol	Failure mode	Failure rate/h FIT
IN	Input channel failed	700
OUT	Output channel failed	31.5
L	Re-triggering line failed	78

Table A.7: Retriggering system failure rates.



Symbol	Failure mode	Failure rate/h FIT
RXD$_1$	RX error detector false alarm	100
RXD$_2$	Missed detection, failed silent	100
IC$_1$	Voter, incorrect result	100
ER$_3$	Energy Conversion, incorrect result	120
VT$_1$	Voter, false Alarm	100
VT$_2$	Missed detection, failed silent	100

Table A.8: BETS failure rates.



Symbol	Failure mode	Failure rate/h FIT
RX$_1$	Optical receiver, invalid format (CRC)	282
RX$_2$	No data, 1ms timeout	400
RX$_3$	Wrong calculated current value	100
RX$_4$	No value calculated (timeout)	100
RXD$_1$	RX error detector, false alarm	100
RXD$_2$	Missed detection	100
VT$_1$	Voter, false Alarm	100
VT$_2$	Missed detection, failed silent	100
WDT$_1$	Watch dog timer, false Alarm	100
WDT$_2$	Missed detection	100
AV$_2$	Average module, incorrect result	100
AV$_3$	No value calculated (timeout)	100
ER$_1$	Energy conversion, incorrect result	175
ER$_2$	No value calculated (timeout)	100
TX$_2$	Transmitter, incorrect coding	100
TX$_3$	Incorrect transmission	100
TX$_4$	No transmission	100

Table A.9: BEMS failure rates.

Symbol	**Failure mode**	**Failure rate/h** **FIT**
REF_1	Incorrect reference voltage	560
AS_1	Multiplexer stuck at an input	62
ADC_1	Conversion, incorrect digital output	1380
AV_1	Averaging, incorrect result	100
TX_1	Incorrect transmission	145
DCCT	General failure	100

Table A.10: BEA failure rates.

Bibliography

[1] J. Arlat, A. Costes, Y. Crouzet, J. Laprie and D. Powell, *Fault Injection and Dependability Evaluation of Fault-Tolerant Systems*, IEEE Transactions on Computers, vol. 42, no. 8, pp. 913-923, Aug. 1993.

[2] J.E. Arsenault and J.A. Roberts (eds.), *Reliability and Maintainability of Electronic Systems*, Computer Science Press, Potomac, 1980.

[3] A. Avizienis, J. Laprie, B. Randell and C. Landwehr, *Basic Concepts and Taxonomy of Dependable and Secure Computing*, IEEE Transactions on Dependable and Secure Computing, vol. 1, No. 1, pp. 11-33, January/March 2004.

[4] A.Vad-Nielsen, *A digital controller for monitoring the trigger and retrigger systems of the LHC beam dump kickers*, CERN thesis 2003.

[5] R. Bell and D. Reinert, *Risk and System Integrity Concepts for Safety-Related Control Systems*, in F. Redmill and T. Anderson, editors, Safety-Critical Systems: Current Issues, Techniques and Standards, pp. 16-42, Chapman & Hall, 1990.

[6] A. Bobbio and K.S. Trivedi, *An Aggregation Technique for the Transient Analysis of Stiff Markov Chains*, IEEE Transactions on Computers, vol. 35, No. 9, pp. 803-814, Sept. 1986.

[7] P.J. Boland, *A Reliability Comparison of Basic Systems Hazard Rate Functions*, Applied Stochastic Models and Data Analysis, vol. 13, pp. 377-384, 1998.

[8] J. Bonthond, J.H. Dieperink, L. Ducimetière, U. Jansson, E.B. Vossenberg, *Dual Branch High voltage Pulse Generator for the Beam Extraction of the Large Hadron Collider*, 25th International Power Modulator Symposium, Hollywood, California, June 2002.

[9] F. Bordry, *LHC Power Converters: performance and requirements*, Proceedings of the LHC Workshop Chamonix XI, Chamonix, France, 15 - 19 January 2001.

[10] M. Bouissou, *Boolean Logic Driven Markov processes: A Powerful New Formalism for Specyfing and Solving Very Large Markov Models*, Proceedings of the 6th International Conference on Probabilistic Safety Assessment and Management, (San Juan, Puerto Rico, USA), 23-28 June 2002.

[11] M. Bouissou and J. L. Bon, *A new formalism that combines advantages of fault-trees and Markov models: Boolean logic Driven Markov Processes*, Reliability Engineering and System Safety, vol. 82, pp. 149-163, Elsevier Science, Nov. 2003.

[12] J. Bowles, *A Survey of Reliability prediction Procedures for Microelectronics Devices*, IEEE Transactions on Reliability, vol. 41, no. 1, pp. 2-12, March 1992.

[13] B. Todd, private communication, CERN 2005.

[14] J. Bukowsky, *Modeling and Analyzing the Effects of Periodic Inspections on the Performance of Safety Critical Systems*, IEEE Transactions on Reliability, vol. 50, no. 3, pp. 321-329, 2001.

[15] E. Carlier and others., *Information exchange between beam dump system and other systems*, Proceedings of the LHC Workshop Chamonix XI, Chamonix, France, 15 - 19 January 2001.

[16] E. Carlier and others, *Design aspects related to the reliability of the control architecture of the LHC beam dump kicker systems*, 9th International

Conference on Accelerator and Large Experimental Physics Control Systems ICALEPCS 2003, Gyeongju, Korea, 13 - 17 October 2003.

[17] E. Carlier and others, *Design aspects related to the reliability of the LHC beam dump kickers systems*, 17th Particle Accelerator Conference PAC '97, Vancouver, Canada, 12 - 16 May 1997, pages 225-227.

[18] C.G. Cassandras and S. Lafortune, *Introduction to Discrete Events Systems*, Kluwer Academic Publisher, 1999.

[19] CERN Documents Server, `http://cds.cern.ch/`, CERN, Ginevra.

[20] H. Choi, V.G. Kulkarni and K.S. Trivedi, *Transient Analysis of Deterministic and Stochastic Petri Nets, Lecture Notes in Computer Science*, vol. 691, M. Ajmone Marsan (ed.), Proc. 14th International Conference on Application and Theory of Petri Nets, Springer-Verlag, Heidelberg, pp. 166-185, 1993.

[21] H. Choi, V.G. Kulkarni and K.S. Trivedi, *Markov Regenerative Stochastic Petri Nets*, Performance Evaluation, vol. 20, no. 1-3, pp. 337-357, 1994.

[22] C.Y. Choi, B.U. Johnson and J.A. Profeta III, *Safety Issues in the Comparative Analysis of Dependable Architectures*, IEEE Transactions on Reliability, vol. 46, no. 3, pp. 393-401, 1997.

[23] E. Ciapala, F. Rodriguez-Mateos, R. Schmidt and J. Wenninger, *The LHC Post-Mortem System*, LHC Project Note 303, CERN, Geneva, 2002.

[24] G. Ciardo, R. Marie, B. Sericola and K.S. Trivedi, *Performability Analysis Using semi-Markov Reward Processes*, IEEE Transactions on Computers, vol. 39, no. 10, pp. 1251-1264, October 1990.

[25] G. Ciardo, R. German and C. Lindemann, *A Characterization of the Stochastic Process Underlying a Stochastic Petri net*, IEEE Transactions on Software Engineering, vol. 20, no. 7, July 1994.

[26] E.G. Coffman and E.N. Gilbert, *Optimal Strategies for Scheduling Checkpoints and Preventive Maintenance*, IEEE Transactions on Reliability, vol.39, no. 1, pp. 9-18, April 1990.

[27] B. Dehning, *Beam loss monitor system for machine protection*, 7th European Workshop on Beam Diagnostics and Instrumentation for Particle Accelerators, DIPAC'05 , Lyon, France, 06 - 08 June 2005.

[28] W. Denson, *The History of Reliability Prediction*, IEEE Transactions on Reliability, vol. 47, no. 3, pp. 321-328, Sept. 1998.

[29] J.B. Dugan, S.J. Bavuso and M.A. Boyd, *Dynamic Fault-Tree Models for Fault-Tolerant Computer Systems*, IEEE Transactions on Reliability, vol. 41, no. 3, pp. 363-377, Sept. 1992.

[30] M. Economou, *The Merits and Limitations of Reliability Predictions*, Reliability Availability Maintainability Symposium, RAMS, pp.352-357, Los Angeles, Jan. 2004.

[31] Reliability Analysis Center, *Electronic Part Reliability Data*, RAC, Rome (NY) USA, 1997.

[32] R. Filippini and A. Bondavalli, *Modeling and Analysis of a Scheduled maintenance System: a DSPN approach*, the Oxford Computer Journal, vol. 47, no. 6, pp. 634-650, Nov. 2004.

[33] R. Filippini, B. Dehning, G. Guaglio, F. Rodriguez-Mateos, R. Schmidt, B. Todd, J. Uythoven, A. Vergara-Fernandez, M. Zerlauth, *Reliability Assessment of the LHC Machine Protection System*, Particle Accelerators Conference PAC 2005, pp. 1257-1259, Knoxville, USA, 16-20 May 2005.

[34] Reliability Analysis Center RAC, *Failure Mode/Mechanism Distributions*, FMD-97, Rome (NY), USA, 1997.

[35] R. German, D. Logothetis, and K. S. Trivedi, *Transient Analysis of Markov Regenerative Stochastics Petri Nets: a comparison of approaches*, Proc. Sixth International Workshop on Petri Nets and Performance Models, pp. 103-112, Durham, 1995.

[36] W.M. Goble and J.Bukowsky, *Defining Mean Time To Failure in a Particular Failure State for Multi-States Systems*, IEEE Transactions on Reliability, vol. 50, no. 2, pp. 221-228, 2001.

[37] G. Guaglio, B. Dehning, *Reliability of Beam Loss Monitors System for the Large Hadron Collider*, CERN PhD thesis, Geneva 2005.

[38] B. R. Haverkort and A. Meeuwissen, *Uncertainty Analysis of Markov Reward Models*, IEEE Transactions on Reliability, vol. 44, no. 1, pp. 147-154, 1995.

[39] A. Hoyland and M. Rausand, *System Reliability Theory: Models and Statistical methods*, Wiley, New York, 1994.

[40] D. Huwe Jones, contribution to the failure modes analysis of the BEM and BET systems of the LBDS, CERN summer student, 2005.

[41] International Atomic Energy Agency, *Protection System and Related Features in Nuclear Power Plants, A Safety Guide*, IAEA Safety Series No 50-SG-D3, Vienna, 1980.

[42] International Electro-technical Commission IEC, *Functional Safety of Electrical-Electronic-Programmable Electronic Safety Related Systems*, IEC 61508 International Standard, Geneva, 1998.

[43] International Electro-technical Commission IEC, *Universal Model for Reliability Prediction of Electronics Components, PCBs and Equipment*, Reliability data Handbook IEC 62380, Geneva, 2004.

[44] The IEEE/PES Task Force, *The present Status of Maintenance Strategies and the Impact of Maintenance on Reliability*, Risk and Probability Applications Subcommittee, IEEE Transactions on Power Systems, vol. 16, no. 4, Nov. 2001.

[45] Institute of Electrical and Electronic Engineers, *IEEE Guide for General principles of Reliability Analysis of Nuclear Power Generating Station Safety Systems*,IEEE 325, New York, 1987.

[46] U. Isaksen, J.P. Bowen, and N. Nissanke, *System and Software Safety in Critical Systems*, The University of Reading, Whiteknights, United Kingdom, December 1996.

[47] B.W. Johnson and Y. Yangyaing, *A Comparison of Two Safety Critical Architectures Using the Safety Related Metrics*, Annual Reliability Availability and Maintainability Symposium, pp. 621-627, Los Angeles, 2004.

[48] J. Jones and J. Hayes, *A Comparison of Electronic-Reliability Prediction Models*, IEEE Transactions on Reliability, vol.48, no. 2, pp. 127-134, June 1999.

[49] M. Kaaniche, J.P. Laprie and J.P. Blanquart, *Dependability Engineering of Complex Computing Systems*, 6th IEEE International Conference on Engineering of Complex Computer Systems (ICECCS'2000), Tokyo, pp. 36-46, Sept. 2000.

[50] P. King and R. Pooley, *Using UML to Derive Stochastic Petri Net Models*, Proceedings of the 50th Annual UK Performance Engineering Workshop, pp. 45-56, July 1999.

[51] M. Lanus, L. Yin and K.S. Trivedi, *Hierarchical Composition and Aggregation of State-Based Availability and Performability Models*, IEEE Transaction on Reliability, vol. 52, no. 1, pp. 44-52, March 2003.

[52] J. Laprie, J. Arlat, C. Beounes and K. Kanoun, *Definition and Analysis of Hardware and Software Fault-Tolerant Architectures*, IEEE Computer,pp 39-51, July 1990.

[53] J. Laprie, *Dependability: Its Attributes, Impairments and Means*, Predictably Dependable Computing Systems, B. Randell et al., eds., pp. 3-24, 1995.

[54] G. Latif-Shabgahi, J.M. Bass and S. Bennet, *A Taxonomy for Software Voting Algorithms Used in Safety Critical Systems*, IEEE Transactions on Reliability, vol.53, no. 3, pp. 319-328, September 2004.

[55] M. Buhler-Broglin, D. Lajust and R. Lewis, *The LEP in short*, CERN 1989.

[56] The LHC Design Report: *The LHC Main Ring*, vol. 1, CERN, Geneva 2004.

[57] The LHC Design Report: *The LHC Injector Chain*, vol. 3, CERN, Geneva 2004.

[58] N. Limnios, *Dependability Analysis of Semi-Markov Systems*, Elsevier Science, Reliability Engineering and System Safety, vol. 55, pp. 203-207, 1997.

[59] D. Logothesis, A. Puliafito and K.S. Trivedi, *Markov Regenerative Models*, Proc. Int. Computer Performance and Dependability Symposium, Erlangen, Germany, pp. 134-143, 1995.

[60] V. Mainkar, H. Choi and K.S. Trivedi, *Sensitivity Analysis of Markov Regenerative Stochastic Petri Nets*, Proc. Fifth Int. Workshop on Petri Nets and Performance Models (PNPM93), Toulouse, 1993.

[61] M. Malhotra, A. Reibman, *Selecting and Implementing Phase Approximations for Semi-Markov Models*, Communication Statistics, Stochastic Models, vol. 9, no. 4, pp. 473-506, 1993.

[62] M. Malhotra and K. S. Trivedi, *Power-hierarchy of dependability model types*, IEEE Transactions on Reliability, vol. 43, pp. 493-502, Sept. 1994.

[63] J. McDermid, *Issues in the Development of Safety Critical Systems*, in F. Redmill and T. Anderson, editors, Safety-Critical Systems: Current Issues, Techniques and Standards, pp. 16-42, Chapman & Hall, 1990.

[64] D.P. McGinnis, *FNAL Tevatron Operational Status*, Particle Accelerator Conference PAC 2005, pp. 484-488, Knoxville, TN, USA, 16 - 20 May 2005.

[65] Military Standard 1629A, *Procedure for Performing a Failure Modes Effects and Criticalities Analysis: Revision A*, Department of Defense, Washington D.C., 1998.

[66] Military Handbook 217F, *Reliability Prediction of Electronic Equipment*, Department of Defense, Washington D.C., 1993.

[67] Military Handbook 338B, *Electronic Reliability Design Handbook: Revision B*, Department of Defense, Washington D.C., 1998.

[68] J. Muppala, R. Fricks, and K. S. Trivedi, *Techniques for System Dependability Evaluation*, in Computational Probability, W. Grassman (ed.), pp. 445-480, Kluwer Academic Publishers, The Netherlands, 2000.

[69] I. Mura and A. Bondavalli, *Markov Regenerative Stochastic Petri Nets to Model and Evaluate Phased Mission Systems Dependability*, IEEE Transactions on Computers, vol. 50, no. 12, pp.1337-1351, 2001.

[70] M. G. Pecht and Wen-Chang Kang,*A Critique of the MIL-HDBK-217E Reliability Prediction Methods*, IEEE Transactions on Reliability, vol. 37, no. 5, pp. 453-457, Dec. 1988.

[71] M. G. Pecht and F.R. Nash, *Predicting the Reliability of Electronic Equipment*, Proceeding of the IEEE, vol. 82, no. 7, pp. 992-1004, July 1994.

[72] N. Phinney and others, *Reliability Simulations for a Linear Collider*, European Particle Accelerator Conference, EPAC 2004, pp.457-459, Lucerne, Switzerland, 7-9 July 2004.

[73] D. Powell, *Failure Mode Assumptions and Assumption Coverage*, Proceeding of the 22th Annual International Symposium on Fault Tolerant Computing FTCS 22, pp. 386-395, IEEE, Boston, MA, USA, July 1992.

[74] D. K. Pradhan, *Fault Tolerant Computer System Design*, Prentice Hall, New Jersey, 1995.

[75] B. Puccio, *Interlock Channels and Their Timescales*, Proceedings of LHC Chamonix Workshop, March 2003.

[76] A.V. Ramesh and K.S. Trivedi, *On the Sensitivity of Transient Solutions of Markov Models*, Proc. 1993 ACM SIGMETRICS Conference, Santa Clara, CA, pp. 122-134, May 1993.

[77] M. Rampl, *Study for a failsafe trigger generation system for the Large Hadron Collider beam dump kicker magnets*, CERN thesis, Geneva, 1999.

[78] A. Reibman and K.S. Trivedi, *Numerical Transient Analysis of Markov Models*, Computers and Operations Research, vol. 15, no. 1, pp. 19-36, 1988.

[79] A. Rindos, S. Woolet, I. Viniotis and K.S. Trivedi, *Exact methods for the Transient Analysis of Nonhomogeneous Continuous Time markov Chains*, 2nd International Workshop on the Numerical Solution of Markov Chains, W. J. Stewart (ed.), Kluwer Academic Publishers, 1995.

[80] S.M. Ross, *Introduction to Probability Models*, Academic Press, London, 2000.

[81] Aad. P. A. van Moorsel and W. H. Sanders, *Transient Solution of Markov Models by Combining Adaptive and Standard Uniformization*, IEEE Transactions on Reliability, vol. 46, no. 3, pp. 430-440, Sept. 1997.

[82] M.L. Shooman, *Reliability of Computer Systems and Networks, Fault Tolerance Analysis and Design*, J.Wiley and Sons, New York 2002.

[83] M.L. Shooman, *Probabilistic Reliability: an Engineering Approach*, Mc-Graw Hill, 1968.

[84] D.P. Siewiorek, *Architecture of Fault Tolerant Computers: A Historical Perspective*, Proceeding of the IEEE, vol. 79, N0. 12, pp. 1710-1734, Dec. 1991.

[85] D.P. Siewiorek and R. Schwarz, *Reliable Computer Systems: Design and Evaluation*, Digital Press, Bedford, Mass., 1992.

[86] D. Todd Smith, T.A. Delong, B.W. Johnson and T.D. Giros, *Determining the Expected Time To Unsafe Failure*, Fifth IEEE International Symposium on High Assurance System Engineering, pp. 17-24, Albuquerque, Nov. 2000.

[87] K.S. Trivedi, *Probability and Statistics with Reliability, Queuing and Computer Science applications*, Prentice-Hall, New Jork, 1982.

[88] K.S. Trivedi, G. Ciardo, M. Malhotra, R. Sahner, *Dependability and Performability Analysis, Performance Evaluation of Computer and Communication Systems*, Lecture Notes in Computer Science, L. Donatiella, R. Nelson (eds.), pp. 587-612, Springer-Verlag, 1993.

[89] J. Uythoven, R. Filippini, B. Goddard, M. Gyr, V. Kain, R. Schmidt, J. Wenninger, *Possible Causes and Consequences of Serious Failures of the LHC Machine Protection System*, 9th European Particle Accelerator Conference, pp. 620-622, EPAC 2004, Lucerne, Switzerland, 5-9 July 2004.

[90] A. Vergara-Fernandez, F. R. Rodriguez-Mateos, *Reliability of the Quench Protection System for the LHC Superconducting Elements*, CERN PhD thesis, Geneva 2003.

[91] G. Watson, *MIL Reliability: a New Approach*, IEEE Spectrum, pp. 46-49, August 1992.

[92] J. Wenninger, R. Schmidt, *Protection Against Accidental Beam Losses at the LHC*, Particle Accelerators Conference PAC 2005, pp. 492-494, Knoxville, USA, 16-20 May 2005.

[93] F. Willeke, *HERA status and perspectives of future lepton hadron colliders*, pp. 51-55, EPAC 2002, Paris, France.

[94] S. Wolfram, *The Mathematica Book*, Fifth Edition, Wolfram Research inc. 2003.

[95] L. Yin, M. Smith and K.S. Trivedi, *Uncertainty Analysis in Reliability Modeling*, Proceedings of the Annual Reliability and Maintainability Symposium, RAMS 2001, Philadelphia, PA, USA, January, 2001.

[96] M. Zerlauth, B. Goddard, V. Kain, R. Schmidt, *Detecting Failures in Electrical Circuits Leading to Very Fast Beams Losses in the LHC*, 9th European Particle Accelerator Conference EPAC 2004 , pp. 1930-1932, Lucerne, Switzerland, 05 - 09 Jul 2004.

[97] M. Zerlauth, private communication, CERN 2005.