

Appendix D

Fault tree analysis

D.1 Introduction

A fault tree is a logic diagram that displays the relationships between a potential critical event (accident) in a system and the reasons for this event. The reasons may be environmental conditions, human errors, normal events (events which are expected to occur during the life span of the system) and specific component failures. A properly constructed fault tree provides a good illustration of the various combinations of failures and other events which can lead to a specified critical event. The fault tree is easy to explain to engineers without prior experience of fault tree analysis.

An advantage with a fault tree analysis is that the analyst is forced to understand the failure possibilities of the system, to a detailed level. A lot of system weaknesses may thus be revealed and corrected during the fault tree construction.

A fault tree is a *static* picture of the combinations of failures and events which can cause the TOP event to occur. Fault tree analysis is thus not a suitable technique for analysing dynamic systems, like switching systems, phased mission systems and systems subject to complex maintenance strategies.

A fault tree analysis may be qualitative, quantitative or both, depending on the objectives of the analysis. Possible results from the analysis may e.g. be:

1. A listing of the possible combinations of environmental factors, human errors, normal events and component failures that can result in a critical event in the system.
2. The probability that the critical event will occur during a specified time interval.

Figure ?? shows an example fault tree for the bike.

The analysis of a system by the fault tree technique is normally carried out in five steps:

1. Definition of the problem and the boundary conditions.

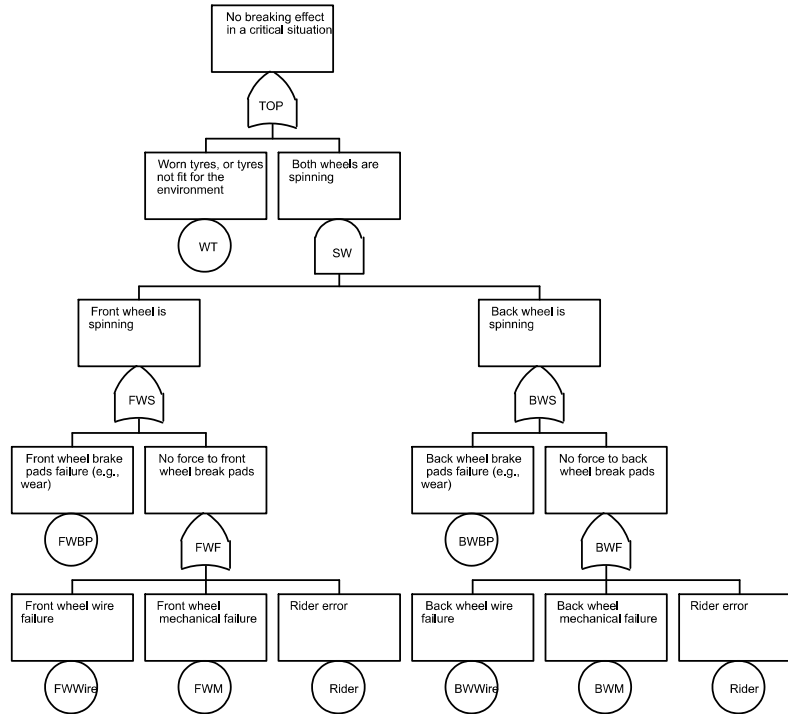


Figure D.1: FTA example for a bike

2. Construction of the fault tree.
3. Identification of minimal cut and/or path sets.
4. Qualitative analysis of the fault tree.
5. Quantitative analysis of the fault tree.

In the following we will present the basic elements of standard fault tree analysis. Then we will conclude this chapter by presenting a numerical example illustrating how the technique could be utilised in relation to maintenance optimisation.

D.2 Fault tree construction

D.2.1 Fault tree diagram, symbols and logic

A fault tree is a logic diagram that displays the connections between a potential system failure (TOP event) and the reasons for this event. The reasons (Basic events) may be environmental conditions, human errors, normal events and component failures. The graphical symbols used to illustrate these connections are called “logic gates”. The output from a logic gate is determined by the input events.

The graphical layout of the fault tree symbols are dependent on what standard we choose to follow.

D.2.2 Definition of the Problem and the Boundary Conditions

This activity consists of:

1. Definition of the critical event (the accident) to be analysed.
2. Definition of the boundary conditions for the analysis.

The critical event (accident) to be analysed is normally called the TOP event. It is very important that the TOP event is given a clear and unambiguous definition. If not, the analysis will often be of limited value. As an example, the event description “Fire in the plant” is far too general and vague. The description of the TOP event should always answer the questions: **What, where and when?**

What: Describes what type of critical event (accident) is occurring, e.g., collision between two trains.

Where: Describes where the critical event occurs, e.g., on a single track section.

When: Describes when the critical event occurs, e.g., during normal operation.

A more precise TOP event description is thus: “Collision between two trains on a single track section during normal operation”.

1. To get a consistent analysis, it is important that the *boundary conditions* for the analysis are carefully defined. By boundary conditions we mean: **The physical boundaries of the system.** What parts of the system are to be included in the analysis, and what parts are not?
2. **The initial conditions.** What is the operational state of the system when the TOP event is occurring? Is the system running on full/reduced capacity? Which valves are open/closed, which pumps are functioning etc.?
3. **Boundary conditions with respect to external stresses.** What type of external stresses should be included in the analysis? By external stresses we here mean stresses from war, sabotage, earthquake, lightning etc.
4. **The level of resolution.** How far down in detail should we go to identify potential reasons for a failed state? Should we as an example be satisfied when we have identified the reason to be a “valve failure”, or should we break it further down to failures in the valve housing, valve stem, actuator etc.? When determining the required level of resolution, we should remember that the detail in the fault tree should be comparable to the detail of the information available

D.2.3 Construction of the Fault Tree

The fault tree construction always starts with the TOP event. We must thereafter carefully try to identify all fault events which are the immediate, necessary and sufficient causes that result in the TOP event. These causes are connected to the TOP event via a logic gate. It is important that the first level of causes under the TOP event is developed in a structured way. This first level is often referred to as the TOP structure of the fault tree. The TOP structure causes are often taken to be failures in the prime modules of the system, or in the prime functions of the system. We then proceed, level by level, until all fault events have been developed to the required level of resolution. The analysis is in other words deductive and is carried out by repeated asking “What are the reasons for...?”

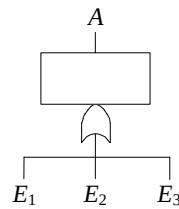


Figure D.2: OR-gate

Figure ?? shows the OR-gate indicating that the output event A occurs if any of the input events E_i occurs. In relation to the bike example with TOP event “No breaking effects” the two events: “No friction” and “both wheels spinning” are connected by an OR gate since any of these events will lead to the TOP event.

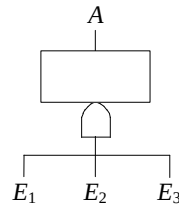


Figure D.3: AND-gate

Figure ?? shows the AND-gate indicating that the output event A occurs only when all the input events E_i occurs simultaneously. In the bike example, “Front wheel is spinning” and “Rear wheel is spinning” are connected by an AND gate, since both these event have to occur in order to full fill the requirement that both wheels are spinning.

Figure ?? shows the Basic event representing a basic equipment fault or failure that requires no further development into more basic faults or failures. An example of a basic event in the bike example is “Breakage in break wire”.

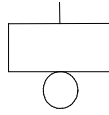


Figure D.4: BASIC-event

D.3 Identification of Minimal Cut- and Path Sets

A fault tree provides valuable information about possible combinations of fault events which can result in a critical failure (TOP event) of the system. Such a combination of fault events is called a cut set.

*A **cut set** in a fault tree is a set of Basic events whose (simultaneous) occurrence ensures that the TOP event occurs. A cut set is said to be **minimal** if the set cannot be reduced without losing its status as a cut set.*

*A **path set** in a fault tree is a set of Basic events whose non-occurrence (simultaneously) ensures that the TOP event does not occur. A path set is said to be **minimal** if the set cannot be reduced without losing its status as a path set.*

In practice only minimal cut sets are used for evaluation of fault trees. To find the minimal cut sets we apply the MOCUS algorithm (Method Of obtaining Cut Sets). The MOCUS algorithm essentially contains the following elements:

1. Start with the TOP event
2. As the algorithm proceeds, the result is stored in a matrix like format of rows and columns
3. AND- and OR-gates are resolved by replacing the gate with its “children” in the fault tree diagram
4. An AND-gate means that the gate is replaced by new elements for the row(s) it is found
5. An OR-gate means that the gate is replaced by as many rows that the gate has children, where each child is inserted at the position of the OR-gate being replaced
6. When all gates are replaced, we remain with only the basic events, where each row corresponds to a cut set

Note that the the cut sets will not necessarily be be minimal. To make the cut sets minimal we have to:

1. Replace duplicates of one event with only one occurrence of that event in each row
2. If one row is a sub set of another row, then the larger of these two rows (representing non-minimal cut sets) is removed

The MOCUS algorithm is demonstrated here: <http://folk.ntnu.no/jvatn/eLearning/TPK4120/Examples/MOCUS.html> in relation to the example used in the lectures.

D.3.1 *koon* gate

The *koon* gate is something “between” the AND and OR gate. A *koon* gate occurs if k out of the n inputs occur. Note that in FTA we focus on fault states, i.e., an event occurring means a failure, hence the “voting” in FTA is different from in RBD. To clarify, the following notation is often used:

- $koon : G$ is used if we consider the functioning of components (G =Good). The system (block) functions if k or more out of the n components are functioning
- $koon : F$ is used if we consider the fault of components (F =Fault state). The system (gate/-TOP event) occurs if k or more out of the n inputs are occurring (i.e., in a fault state)

Note the following relation:

$$koon : G = (n - k + 1)oon : F$$

$$koon : F = (n - k + 1)oon : G$$

Consider a system with three pumps each having 50% capacity. The system functions if at least 2 of the pumps are functioning. In an RBD we then use the $2oo3 : G$ block for this system, and for the FTA we use the $koon : F = n - k + 1oon : G = 3 - 2 + 1oo3 = 2oo3$ gate.

If we have 4 such pumps, the RBD representation is $2oo4 : G$, and in FTA we use the $koon : F = n - k + 1oon : G = 4 - 2 + 1oo4 = 3oo4$ gate meaning that 3 or more pumps must be in a fault state in order to give a system failure (TOP event).

Computerized FTA programs will offer the $koon : F$ as part of the drawing palette. For manual construction of a fault tree with a $koon : F$ gate we can use an OR-gate followed by several AND-gates. Each AND-gate is then a sub-set with k out of the n inputs. There are altogether $\binom{n}{k}$ ways we may choose k inputs out of n inputs, hence we will have $\binom{n}{k}$ AND-gates to put under the OR-gate.

D.4 Qualitative Evaluation of the Fault Tree

A qualitative evaluation of the fault tree may be carried out on the basis of the minimal cut sets. The importance of a cut set depends obviously on the number of Basic events in the cut set. The number of different Basic events in a minimal cut set is called the *order* of the cut set. A cut set of order one is usually more critical than a cut set of order two, or higher. When we have a cut

set with only one Basic event, the TOP event will occur as soon as this Basic event occurs. When a cut set has two Basic events, both of these have to occur at the same time to cause the TOP event to occur.

Another important factor is the type of Basic events in a minimal cut set. We may rank the criticality of the various cut sets according to the following ranking of the Basic events:

1. Human error
2. Failure of active equipment
3. Failure of passive equipment

The ranking is based on the assumption that human errors occur more frequently than active equipment failures, and that active equipment is more failure-prone than passive equipment (an active or running pump is for example more exposed to failures than a passive standby pump).

D.5 Quantitative analysis

In the quantitative part of a fault tree analysis the main objective is to calculate the following metrics:

- $Q_0(t)$ = Probability that the TOP-event occurs at time t
- $F_0(t)$ = Expected number of TOP-event occurrence per unit time at time t
- $I(i | t)$ = Importance metric for basic event i at time t

For the calculations we need the minimal cut set as well as basic event frequencies and probabilities.

D.5.1 Upper Bound Approximation, $Q_0(t)$

Assume that we have found the minimal cut sets of the fault tree, i.e., K_j . Further assume that the minimal cut sets do not contain common components, hence they are independent (also provided that the components are independent). We may now arrange the cut set in a series structure as indicated in Figure ???: Let E_j denote the event that cut set number j is occurring. The probability that cut set number j is occurring is found by:

$$\Pr(E_j) = \check{Q}_j(t) = \prod_{i \in K_j} q_i(t)$$

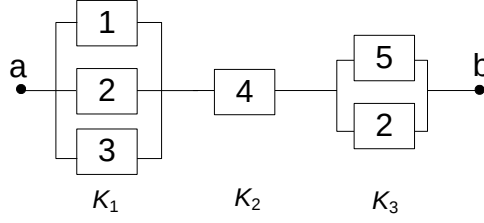


Figure D.5: Example cut set structure

We now have

$$\begin{aligned} Q_0(t) &= \Pr(\text{TOP event occurs at time } t) = 1 - \Pr(\text{TOP event does not occur at time } t) \\ &= 1 - \Pr(\text{No cut set occurs at time } t) \end{aligned}$$

Since the cut sets are independent, and the probability that cut set number j is occurring is given by $\check{Q}_j(t)$, we have:

$$Q_0(t) = 1 - \prod_{j=1}^k (1 - \check{Q}_j(t))$$

where

$$\check{Q}_j(t) = \prod_{i \in K_j} q_i(t)$$

Generally there might be some basic events that occur in two or more cut sets, hence the cut sets are *dependent*, and it may be proven that the formula represents an upper bound for the TOP event probability:

$$Q_0(t) \leq 1 - \prod_{j=1}^k (1 - \check{Q}_j(t))$$

Hence, we may use:

$$Q_0(t) \approx 1 - \prod_{j=1}^k (1 - \check{Q}_j(t))$$

which is referred to as the upper bound approximation and is usually considered to be a good approximation when the $q_i(t)$ s are small.

To argue for the less or equal sign we realize that cut sets are “positive dependent” if they

have common components. For two cut sets we have

$$\Pr(E_1^C \cap E_2^C) = \Pr(E_1^C | E_2^C) \Pr(E_2^C) > \Pr(E_1^C) \Pr(E_2^C)$$

and

$$Q_0 = 1 - \Pr(E_1^C \cap E_2^C) < 1 - \Pr(E_1^C) \Pr(E_2^C) = 1 - (1 - \check{Q}_1)(1 - \check{Q}_2)$$

and we may give similar arguments for more two or more cut sets.

D.5.2 The Inclusion-Exclusion Principle, $Q_0(t)$

Referring to Figure ?? it is also obvious that we may write:

$$Q_0(t) = \Pr(\cup_j E_j)$$

A challenge here is to find the probability of the union of events. For two events A and B we have $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$. For more than two events (cut sets) this becomes more complicated, and we have to use the general addition theorem in probability:

$$Q_0(t) = \Pr(\cup_j E_j) = \sum_j \Pr(E_j) - \sum_{i < j} \Pr(E_i \cap E_j) + \sum_{i < j < k} \Pr(E_i \cap E_j \cap E_k) - \dots$$

To find $\Pr(E_i \cap E_j)$, $\Pr(E_i \cap E_j \cap E_k)$ is straight forward since these intersections of events are in fact intersection of a set of basic events, and we may multiply the corresponding probabilities as we have done for a single minimal cut set. The challenge is the number of terms we have to calculate. As a starting point we can only take the first sum, i.e., adding the cut set occurrences for each cut set. A slightly better approach would be to subtract the next sum. There are some ways we can optimize the calculations, and finding bounds for the answer to use as a stopping rule, see the textbook. Very often the inclusion-exclusion principle is used by only adding the cut set probabilities:

$$Q_0(t) \approx \sum_{j=1}^k \check{Q}_j(t) \tag{D.1}$$

which is faster than the upper bound approximation, but less accurate.

The next challenge is to find the basic event probabilities, $q_i(t)$. Three situations are often considered:

D.5.3 Non-repairable components

If a component cannot be repaired, the probability that it is in a fault state at time t equals $1 - R(t)$, and provided that the component has an exponentially distributed life time, we therefore have:

$$q_i(t) = 1 - e^{-\lambda_i t} \quad (\text{D.2})$$

where λ_i is the constant failure rate of the component.

D.5.4 Repairable components

To derive $q_i(t)$ for a repairable components we may use Markov analysis. The probability that the component is in a fault state at time t is then shown to be (according to eq. 8.22):

$$q_i(t) = \frac{\lambda_i}{\mu_i + \lambda_i} \left(1 - e^{-(\lambda_i + \mu_i)t} \right) \quad (\text{D.3})$$

where λ_i is the constant failure rate of the component, and $\mu_i = 1/\text{MDT}_i$ is the constant repair rate. When t is large compared to $\frac{1}{\lambda_i + \mu_i}$ we have

$$q_i(t) \approx \frac{\lambda_i}{\mu_i + \lambda_i} \approx \lambda_i \text{MDT}_i \quad (\text{D.4})$$

if repair times are short compared to failure times. If this holds, it is safe to use this approximation when $t > 3\text{MDT}_i$, where MDT_i is the mean time to restoration for the component.

D.5.5 Periodically tested components

For components with a hidden function, it is usual to perform a functional test at fixed time intervals, say τ_i , to verify that the component is able to carry out it's function. It may be shown that the (on demand) failure probability of such a component is given by:

$$q_i(t) \approx \lambda_i \tau_i / 2 \quad (\text{D.5})$$

q_i is often referred to as the probability of failure on demand (PFD).

D.5.6 TOP event frequency, $F_0(t)$

$F_0(t)$ denotes the expected number of occurrences of the TOP event per unit time. In principle we may calculate $F_0(t)$ at various point of times, but usually we focus on the steady state situation, and therefore we omit the time dependency, i.e., we seek F_0 .

The arguments are as follows:

- We know the minimal cut sets
- If one cut set should be the “contributor” to the TOP event to occur, the other cut sets cannot be occurring
- For a basic event in one cut set to bring the cut set to occur, requires that all other basic events in that cut set are occurring

Let $C_{\mathcal{K}}$ denote a minimal cut set, then the cut set occurrence frequency is given by:

$$\check{w}_{\mathcal{K}} = \sum_{i \in C_{\mathcal{K}}} w_i \prod_{\ell \in C_{\mathcal{K}}, \ell \neq i} q_{\ell} \quad (\text{D.6})$$

where w_i is the ROCOF of basic event i , and q_l is the probability that basic event l is occurring.

The ROCOF is the rate of occurrence of failures. To define the ROCOF we need to have a stochastic process perspective, i.e., we consider what is happening in a time interval rather than things are happening in this interval. Let $N(t)$ be the number of failures that occur in $(0, t]$ and let $W(t) = E[N(t)]$. The ROCOF at time t is now defined by

$$w(t) = \lim_{\Delta t \rightarrow 0} \frac{E[N(t + \Delta t) - N(t)]}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{W(t + \Delta t) - W(t)}{\Delta t} = \frac{d}{dt} W(t) \quad (\text{D.7})$$

To obtain the TOP event frequency we may now sum over the $\check{w}_{\mathcal{K}}$'s. However, note that $\check{w}_{\mathcal{K}}$ will not contribute to the TOP event frequency if one of the other cut set is already in a fault state, hence the TOP event frequency is better approximated by:

$$F_0 = w_{\text{TOP}} \approx \sum_{\mathcal{K}=1}^k \check{w}_{\mathcal{K}} \prod_{j=1, j \neq \mathcal{K}}^k (1 - \check{Q}_j) \approx \sum_{\mathcal{K}=1}^k \check{w}_{\mathcal{K}} \frac{1 - Q_0}{1 - \check{Q}_{\mathcal{K}}} \quad (\text{D.8})$$

The formula in Equation (D.8) is the best we can do, but usually $\check{Q}_j$ is rather small, and it will be sufficient to use

$$F_0 \approx \sum_{\mathcal{K}=1}^k \check{w}_{\mathcal{K}} = \sum_{\mathcal{K}=1}^k \sum_{i \in C_{\mathcal{K}}} w_i \prod_{\ell \in C_{\mathcal{K}}, \ell \neq i} q_{\ell} \quad (\text{D.9})$$

The ROCOF of the basic events is usually found by the failure rate, say λ_i . However, a more exact calculation will also take into account the downtime on basic event level, i.e., we may use:

$$w_i = \lambda_i(1 - q_i) \approx \lambda_i \quad (\text{D.10})$$

D.6 Reliability Importance Metrics

In the literature very many reliability importance metrics are presented. We only focus on the following:

- Birnbaum's metric
- Improvement Potential
- The criticality importance metric
- Fussel-Vesley's metric

In principle a metric is linked to basic events. Very often these basic events are component failures, hence the term component importance is often used. There are many reasons to investigate component importance:

- Considering improving the inherent reliability of critical components
- Establish a preventive maintenance program for the most critical components
- Ensure that we have sufficient spare parts for critical components
- Considering implementing (extra) redundancy at component level for the most critical components
- Given that we have a system failure, which component is the most likely to have caused this?

Several measures are discussed, and the various measures will have their strength and weakness to answer the questions above.

D.6.1 Birnbaum's Metric of Reliability Importance

Birnbaum's metric of reliability importance of a component is a sensitivity measure expressing the change in system reliability if component i is slightly changed, i.e.,

$$I^B(i | t) = \frac{\partial Q_0(t)}{\partial q_i(t)} \quad (\text{D.11})$$

It follows that a small change $\Delta p_i(t)$ in the component reliability will result in the following change in system reliability:

$$\Delta Q_0(t) = I^B(i | t) \Delta q_i(t) \quad (\text{D.12})$$

A disadvantage with Birnbaum's metric is that it is difficult to calculate. If we are able to write down the system reliability function, it should be rather easy to find Birnbaum's measure. But in practice we will not be able to write down the TOP event probability, and hence we cannot derive Birnbaum's metric. In some cases we may utilize that:

$$I^B(i | t) = Q_0(t | q_i = 1) - Q_0(t | q_i = 0)$$

It may be shown that $I^B(i | t)$ is the probability that component i is critical at time t . This is a valuable result used in maintenance optimization. Often we need to calculate the expected cost of a failure of a specific component. The contribution to downtime depends on whether the system is down or not, and if a failure will cause a system failure. The Birnbaum's metric is exactly what we need, i.e., we should only include downtime cost if the component under consideration is critical, and $I^B(i | t)$ is then used for calculating this probability.

D.6.2 Improvement Potential

The Improvement Potential states how much the system reliability will increase if component i is replaced with a perfect component:

$$I^{IP}(i | t) = Q_0(t) - Q_0(t | q_i = 0) \quad (D.13)$$

It is easy to show the following relation to Birnbaum's metric:

$$I^{IP}(i | t) = I^B(i | t) q_i(t) \quad (D.14)$$

D.6.3 Criticality Importance

The criticality importance metric $I^{CR}(i | t)$ of component i at time t is the probability that component i is critical for the system and is failed at time t , when we know that the system is failed at time t . It is easy to show the following relation to Birnbaum's metric:

$$I^{CR}(i | t) = \frac{I^B(i | t) \cdot q_i(t)}{Q_0(t)}$$

Fussell-Vesely's Metric

The Fussell-Vesely's importance metric $I^{FV}(i | t)$ of component i at time t is the probability that at least one minimal cut set that contains component i is failed at time t , when we know that the system is failed at time t .

In order to calculate $I^{\text{VF}}(i | t)$ we need some reasoning. We simplify and skip the index t . Now introduce the following notation (we use the terminology “component” whereas the precise word would be “basic event”):

- D_i : At least one minimal cut containing component i is failed
- C : The system is failed
- m_i : Number of minimal cut set containing component i
- E_j^i : Minimal cut set j containing component i is failed

From the definition we have:

$$I^{\text{FV}}(i) = \Pr(D_i | C) = \frac{\Pr(D_i \cap C)}{\Pr(C)} \quad (\text{D.15})$$

Since D_i is a subset of C , then $D_i \cap C = D_i$ and we have:

$$I^{\text{FV}}(i) = \frac{\Pr(D_i)}{\Pr(C)} \quad (\text{D.16})$$

To find $\Pr(D_i)$ we use the same approach as for the “upper bound” approximation for Q_0 . However, note that $D_i = E_1^i \cup E_2^i \cup \dots \cup E_{m_i}^i$ where the union is only taken over minimal cut sets containing component i . This gives:

$$\Pr(D_i) = 1 - \Pr(E_1^{iC} \cap E_2^{iC} \cap \dots \cap E_{m_i}^{iC}) \leq 1 - \Pr(E_1^{iC})\Pr(E_2^{iC})\dots\Pr(E_{m_i}^{iC})$$

$\Pr(E_j^{iC})$ is then obtained by one minus the probability for the event that minimal cut set j is failed, i.e., $\Pr(E_j^{iC}) = 1 - \check{Q}_j^i = 1 - \prod_{l \in K_j} q_l$. The following approximation is usually sufficient to calculate Fussell-Vesely’s measure:

$$I^{\text{FV}}(i) \approx \frac{1 - \prod_j^{m_i} (1 - \check{Q}_j^i)}{Q_0}$$

where the product is over minimal cut sets which contain component i .

If cut set failure probabilities are small, a faster approximation is given by:

$$I^{\text{FV}}(i) \approx \frac{\sum_j^{m_i} \check{Q}_j^i}{Q_0} \quad (\text{D.17})$$

where the sum is over minimal cut sets which contain component i .

By comparing the definition of $I^{\text{CR}}(i)$ and $I^{\text{FV}}(i)$, we see that these measures are rather close to each other. Thus by assuming $I^{\text{CR}}(i) \approx I^{\text{FV}}(i)$, we could easily get an approximation of Birn-

baum's measure from:

$$I^B(i) = \frac{I^{CR}(i) \cdot Q_0}{q_i} \approx \frac{I^{VF}(i) \cdot Q_0}{q_i}$$

D.6.4 System failure frequency obtained by $I^B(i)$

An alternative way to calculate system failure frequency, F_0 , is to start with Birnbaum's measure. First we recall that $I^B(i)$ is the probability that the system is in such a state that component i is critical. That a component is critical means that the system is in such a state that the system is functioning if component i is functioning, and in a fault state if component i is failed. Then it follows that:

$$F_0 = \sum_i I^B(i)(1 - q_i)\lambda_i \quad (\text{D.18})$$

where $p_i = 1 - q_i$ is the probability that component i is functioning, and λ_i is the failure rate of component i . Thus, the contribution of component i to F_0 is given as the product of:

- The probability that component i is critical, i.e., the state of other components
- The probability that component i is functioning
- The failure rate of component i