

## E-Companion of: Application Development Using Fault Data

## Appendix A. Proofs

We prove Theorem 2.1 for the case when  $g(\tau)$  is strictly convex and differentiable through a series of lemmas. The general case follows similarly.

LEMMA A.1 *There exists a unique threshold  $F_{M-1}(t)$  for the  $(M-1)$ st release, and the optimal cost function is given by*

$$V_{M-1}(B, t) = \begin{cases} K_M + K_{M-1} + c_{M-1}B & \text{if } B \geq F_{M-1}(t), \\ + c_M EG(0, T_t^M) & \\ K_M + c_M EG(B, T_t^M) & \text{if } B < F_{M-1}(t). \end{cases} \quad (13)$$

PROOF. Since a final coordination is enforced, the optimal cost function for the last release is

$$V_M(B, t) = K_M + c_M B. \quad (14)$$

Substituting (14) into (7) and (8) with  $m = M - 1$  gives

$$V_{M-1}^c(B, t) = K_M + K_{M-1} + c_{M-1}B + c_M EG(0, T_t^M), \\ V_{M-1}^n(B, t) = K_M + c_M EG(B, T_t^M).$$

Then,

$$V_{M-1}^c(B, t) - V_{M-1}^n(B, t) = K_{M-1} + c_M EG(0, T_t^M) - (c_M EG(B, T_t^M) - c_{M-1}B). \quad (15)$$

By assumption, the right-hand side of (15) is decreasing in  $B$ . Also,

$$V_{M-1}^c(0, t) - V_{M-1}^n(0, t) = K_{M-1} > 0.$$

Hence there is a unique  $F_{M-1}(t)$  such that (15) equals zero at  $B = F_{M-1}(t)$ . This concludes the proof.  $\square$

LEMMA A.2  $V_m^c(B, t) - V_m^n(B, t)$  is strictly decreasing in  $B$ .

PROOF. We prove the result by induction. From the proof of Lemma A.1, the result is true for  $m = M - 1$ . Suppose the result is true for  $m + 1$ . Note that  $V_{m+1}^c(0, t) - V_{m+1}^n(0, t) = K_{m+1} > 0$ . Hence, the induction hypothesis would imply that there is a unique threshold  $F_{m-1}(t)$  such that

$$V_{m+1}(B, t) = \begin{cases} V_{m+1}^n(B, t) & \text{if } B < F_{m+1}(t), \\ V_{m+1}^c(B, t) & \text{if } B \geq F_{m+1}(t). \end{cases}$$

Now we consider the cost function for the  $m$ th release. If coordination is carried out, the optimal cost is

$$V_m^c(B, t) K_m + c_m B + EV_{m+1}[G(0, T_t^{m+1}), t + T_t^{m+1}].$$

If no coordination is carried out, there are two cases to consider. Denote

$$\tilde{F}_{m+1}(t) = \max\{F_{m+1}(t + \tau) : \tau \in S_{T_t^{m+1}}\},$$

where  $S_{T_t^{m+1}}$  is the support of  $T_t^{m+1}$ .

Case 1: If  $B < \tilde{F}_{m+1}(t)$ , then there is a nonempty set  $\mathcal{T} = \{T | G(B, T) < F_{m+1}(t + T)\}$ . Also define  $\tilde{\mathcal{T}} = \{T | G(B, T) \geq F_{m+1}(t + T)\}$ . Thus, the optimal cost function is given by

$$V_m^n(B, t) = \int_{\mathcal{T}} V_{m+1}^n[G(B, T), t + T] d\Phi_t^{m+1}(T) \\ + \int_{\tilde{\mathcal{T}}} V_{m+1}^c[G(B, T), t + T] d\Phi_t^{m+1}(T). \quad (16)$$

By induction hypothesis, we have

$$\frac{\partial V_{m+1}^n(B, t)}{\partial B} > \frac{\partial V_{m+1}^c(B, t)}{\partial B} = c_{m+1}.$$

Hence,

$$\frac{\partial V_m^n(B, t)}{\partial B} = \int_{\mathcal{T}} \frac{\partial V_{m+1}^n[G(B, T), t + T]}{\partial G} \frac{\partial G(B, T)}{\partial B} d\Phi_t^{m+1}(T) \\ + \int_{\tilde{\mathcal{T}}} \frac{\partial V_{m+1}^c[G(B, T), t + T]}{\partial G} \frac{\partial G(B, T)}{\partial B} d\Phi_t^{m+1}(T) \\ > \int_0^\infty c_{m+1} \frac{\partial G(B, T)}{\partial B} d\Phi_t^{m+1}(T) > c_{m+1}.$$

Case 2: If  $B \geq \tilde{F}_{m+1}$ , then the coordination is carried out for sure at the  $(m+1)$ st release. The optimal cost function for not coordinating is

$$V_m^c(B, t) = \int_0^\infty V_{m+1}^c[G(B, T), t + T] d\Phi_t^{m+1}(T). \quad (17)$$

Then

$$\frac{\partial V_m^c(B, t)}{\partial B} = \int_0^\infty c_{m+1} \frac{\partial G(B, T)}{\partial B} d\Phi_t^{m+1}(T) > c_{m+1}.$$

Combine the two cases, we have

$$\frac{\partial V_m^c(B, t)}{\partial B} - \frac{\partial V_m^n(B, t)}{\partial B} < c_m - c_{m+1} < 0.$$

Hence, we conclude the lemma together with the fact that  $V_m^c(0, t) - V_m^n(0, t) = K_m$ .  $\square$

PROOF OF THEOREM 2.1. Follows directly from Lemma A.2.  $\square$

PROOF OF PROPOSITION 2.2. We first note that  $G(B, T) = B + (B + k_c)(e^{\beta T} - 1)$  is linear in  $B$ . Also  $EG(B, T) = B + (B + k_c)(q_t^m - 1)$ . Substituting it into (13), we obtain

$$V_{M-1}(B, t) = \begin{cases} K_M + K_{M-1} + c_{M-1}B & \text{if } B \geq F_{M-1}(t), \\ + c_M(q_t^{M-1} - 1)k_c & \\ K_M + c_M q_t^{M-1}B & \text{if } B < F_{M-1}(t). \end{cases} \quad (18)$$

Since  $q_t^m = Ee^{\beta T_t^{m+1}} > 1$ , we deduce that  $V_{M-1}(B, t)$  is concave in  $B$ . Note also  $F_{M-1}(t) = K_{M-1}/(c_M q_t^{M-1} - c_{M-1})$ .

Now suppose that  $V_{m+1}(B)$  is concave in  $B$ . Then, for  $0 \leq \theta \leq 1$ , we have

$$\begin{aligned} & \theta V_m^n(B_1, t) + (1 - \theta)V_m^n(B_2, t) \\ &= \theta E \left[ V_{m+1} \left( G(B_1, T_t^{m+1}), t + T_t^{m+1} \right) \right] \\ & \quad + (1 - \theta) E \left[ V_{m+1} \left( G(B_2, T_t^{m+1}), t + T_t^{m+1} \right) \right] \\ & \leq E \left[ V_{m+1}(\theta G(B_1, T_t^{m+1}) \right. \\ & \quad \left. + (1 - \theta)G(B_2, T_t^{m+1}), t + T_t^{m+1}) \right] \\ &= V_{m+1}^n \left( \theta B_1 + (1 - \theta)B_2, t + T_t^{m+1} \right). \end{aligned}$$

Since  $V_m^c(B, t)$  is linear, we conclude that  $V_m(B, t)$  is concave in  $B$ .

To show  $V_m^n(B, t)$  is smooth, we only need to show that  $\partial V_m^n / \partial B|_{B \rightarrow \bar{B}^-} = q_t^m$ . This is established by differentiating (16) with respect to  $B$ .  $\square$

PROOF OF PROPOSITION 2.3. Let  $B^0 = 0$  and  $B^k = \sum_{j=1}^k \alpha_j \delta$  be the  $k$ th break point on  $g(t)$ . Also write  $T = n\delta + \zeta$ , where  $n$  is a nonnegative integer and  $0 \leq \zeta < \delta$ . Define  $B_\delta^k = B^k - \alpha_k \zeta$ . Then

$$\begin{aligned} G(B^k, T) &= B^{k+n} + \alpha_{k+n+1}\zeta \\ G(B_\delta^k, T) &= B^{k+n}. \end{aligned}$$

Note that  $G(B, T)$  is piecewise linear in  $B$  with the break points  $\{B^0, B_\delta^1, B^1, B_\delta^2, \dots\}$ . We examine the

slopes of consecutive pieces of  $G(B, T)$ :

$$\begin{aligned} & \frac{G(B^k, T) - G(B_\delta^k, T)}{B^k - B_\delta^k} \\ &= \frac{B^{k+n} + \alpha_{k+n+1}T - B^{k+n}}{B^k - B^k + \alpha_{k+1}T} \\ &= \frac{\alpha_{k+n+1}}{\alpha_{k+1}} \end{aligned}$$

and

$$\begin{aligned} & \frac{G(B_\delta^k, T) - G(B^{k-1}, T)}{B_\delta^k - B^{k-1}} \\ &= \frac{B^{k+n} - B^{k+n-1} - \alpha_{k+n}T}{B^k - \alpha_k T - B^{k-1}} \\ &= \frac{\alpha_{k+n}}{\alpha_k}. \end{aligned}$$

Clearly, the slopes of  $G(B^k, T)$  is monotone increasing (decreasing) if  $\alpha_{k+1}/\alpha_k$  is monotone increasing (decreasing). Note also that  $T$  is arbitrary. Hence, we conclude the proof.  $\square$

PROOF OF PROPOSITION 2.4. It follows from an inductive argument and the fact that the composition of an increasing concave function to a concave function is concave.  $\square$

PROOF OF THEOREM 2.2. We prove the result for the case when  $G_m^i(B, T)$  is differentiable in  $B$ . At the  $m$ th release, define  $\tau(\bar{B})$  to be the solution of (10).

After  $T_t^{m+1}$  time units, the  $(m+1)$ st module is released, and the effective fault level becomes

$$\bar{G}_{m+1}(\bar{B}, T_t^{m+1}) = \sum_{i=1}^k \alpha_{m+1}^i G^i(0, \tau(\bar{B}) + T_t^{m+1}).$$

The function  $\bar{G}_{m+1}(\bar{B}, T_t^{m+1})$  represents the effective fault level at the  $(m+1)$ st release given that the effective fault level at the  $m$ th release is  $\bar{B}$ . Note that the growth of the effective fault level depends on the release index  $m$  due to its dependence on  $c_m^i$ .

From the proof of Theorem 2.1, the optimality of a threshold policy can be established if we show that

$$\bar{c}_{m+1} E \frac{\partial \bar{G}_{m+1}(\bar{B}, T_t^{m+1})}{\partial \bar{B}} - \bar{c}_m > 0.$$

Differentiating (10) with respect to  $\bar{B}$  and solving for  $d\tau/d\bar{B}$ , we obtain

$$\frac{d\tau}{d\bar{B}} = \frac{1}{\sum_{i=1}^k \alpha_m^i \frac{\partial G^i(0, \tau)}{\partial \tau}} = \frac{\bar{c}_m}{\sum_{i=1}^k c_m^i \frac{\partial G^i(0, \tau)}{\partial \tau}}.$$

Also, since  $G^i(B, T) - B$  increases in  $B$ , it follows directly that  $\partial G^i(0, \tau)/\partial \tau$  increases in  $\tau$ . Hence, we have

$$\begin{aligned} & \bar{c}_{m+1} E \frac{\partial \bar{G}_{m+1}(\bar{B}, T_t^{m+1})}{\partial \bar{B}} \\ &= \bar{c}_{m+1} E \frac{\partial \sum_{i=1}^k \alpha_{m+1}^i G^i(0, \tau(\bar{B}) + T_t^{m+1})}{\partial \bar{B}} \\ &= \sum_{i=1}^k E \frac{\partial c_{m+1}^i G^i(0, s)}{\partial s} \Big|_{s=\tau(\bar{B})+T_t^{m+1}} \\ & \quad \cdot \frac{\bar{c}_m}{\sum_{i=1}^k c_m^i \frac{\partial G^i(0, \tau)}{\partial \tau} \Big|_{\tau=\tau(\bar{B})}} \\ &> \bar{c}_m. \end{aligned}$$

Thus, we conclude the proof.  $\square$

**PROOF OF PROPOSITION 3.1.** From Lemma A.1, the threshold  $F_{M-1}$  satisfies

$$c[EG(B, T) - B] = K + CEG(0, T).$$

At the  $m$ th release, suppose  $F_m > F_{M-1}$ . Then we have

$$\begin{aligned} & V_m^c(F_{M-1}) - V_m^n(F_{M-1}) \\ &= K + cF_{M-1} + V_m(0) \\ & \quad - [K + cEG(F_{M-1}, T) + V_{m+1}(0)] \\ &= V_m(0) - c[EG(F_{M-1}, T) - F_{M-1}] - V_{m+1}(0) \\ &= V_m(0) - [K + cEG(0, T) + V_{m+1}(0)] \\ &\leq 0, \end{aligned}$$

which indicates that coordination should be carried out at  $B = F_{M-1}$ . Thus, we have a contradiction, which proves the lemma.  $\square$

**PROOF OF PROPOSITION 3.2.** In the reverse-time setting, let  $W_l(B) = V_m(B) - V_m(0)$  with  $V_m(\cdot)$  defined in (5), (7) and (8). Then

$$W_{l+1}(B) = \min \{EW_l^n(G(B, T)), K + cB\}, \quad (19)$$

and

$$\begin{aligned} W_l^n(B) &= \int_0^{T_{l-1}(B)} W_{l-1}^n(G(B, T)) d\Phi(T) \\ & \quad - \int_0^{T_{l-1}(0)} W_{l-1}^n(G(0, T)) d\Phi(t) \\ & \quad + \int_{T_{l-1}(B)}^\infty [K + cG(B, T)] d\Phi(T) \\ & \quad - \int_{T_{l-1}(0)}^\infty [K + cG(0, T)] d\Phi(T), \end{aligned}$$

where  $T_l(B)$  solves  $G(B, T_l) = f_l$  and  $f_l$  is the optimal threshold value in Theorem 2.1.

Clearly, if the value functions  $W_l(\cdot)$  converge, then the threshold values must also converge. Thus, we try

to show that  $W_l^n(\cdot)$  converges in  $l$ , which, in turn implies the convergence of  $W_l(\cdot)$ .

Let  $S$  denote the collection of all the value functions  $W_l^n$ ,  $l > 2$ , in the dynamic system. We first note that every element in  $S$  is continuously differentiable. It follows that  $S$  is *equicontinuous*. That is, for any  $\delta > 0$ , there is a  $\varepsilon > 0$ , such that  $|w(x) - w(y)| < \delta$  for  $y \in (x - \varepsilon, x + \varepsilon)$  and each  $w \in S$ . Clearly,  $S$  is *point-wise bounded*. That is,  $w(x) \leq K + cx < \infty$  for each  $x \in [0, \infty)$  and each  $w \in S$ . Thus, by the Arzela-Ascoli theorem (page 245 of Rudin 1987), there is a subsequence of  $\{w_l^n\}$  that converges uniformly as  $k_l \rightarrow \infty$ . Hence, we conclude the proposition.  $\square$

**Proposition A.1** Suppose that the inter-release times are independent. Let  $P_m$ ,  $1 \leq m \leq M$ , be the probability of coordinating at the  $m$ th release. Also, let  $P_0 = 1$ . Then  $P_M = 1$  and

$$P_m = \sum_{i=0}^{m-1} P_{m|i} P_i (1 - P_{i+1|i}) \dots (1 - P_{m-1|i}) \quad (20)$$

$$\text{for } 1 \leq m \leq M - 1, \quad (21)$$

where

$$P_{m|i} = \int_{\frac{1}{\beta} \ln \frac{F_m + kc}{k_c}}^\infty d\Psi(t) \quad (22)$$

and  $\Psi = \Phi^i * \dots * \Phi^m$  is the convolution of  $\Phi^i, \dots, \Phi^m$ .

**PROOF.** Let  $B_m$  be the number of faults at the end of the  $m$ th release. Also denote  $t_m$  as the stopping time to the  $m$ th release. Under the optimal threshold policy, the probability of coordinating at release 1 is given by

$$\begin{aligned} P_1 &= P\{B_1 \geq F_1\} = P\{k_c(e^{\beta t_1} - 1) > F_1\} \\ &= \int_{\frac{1}{\beta} \ln \frac{F_1 + kc}{k_c}}^\infty d\Phi^1(t). \end{aligned} \quad (23)$$

If the last coordination is at the  $i$ th release, then the random variable  $t_m - t_i$ ,  $i < m$ , has the distribution  $\Phi^i * \dots * \Phi^m$ . Thus, the conditional probability of coordinating at the  $m$ th release is given by

$$\begin{aligned} & P_{m|i} \\ &= P\{B_m \geq F_m | B_i \geq F_i; B_k < F_k, i < k < m\} \\ &= P\{k_c(e^{\beta(t_m - t_i)} - 1) \geq F_m | B_i \geq F_i; B_k < F_k, i < k < m\} \\ &= \int_{\frac{1}{\beta} \ln \frac{F_m + kc}{k_c}}^\infty d\Phi^i * \dots * \Phi^m(t). \end{aligned}$$

Using the conditioning rule, we have

$$\begin{aligned}
& P\{B_i \geq F_i; B_k < F_k, i < k < m\} \\
&= P\{B_{m-1} < F_{m-1} | B_i \geq F_i; B_k < F_k, i < k < m-1\} \\
&\quad \cdot P\{B_i \geq F_i; B_k < F_k, i < k < m-1\} \\
&= P\{B_{m-1} < F_{m-1} | B_i \geq F_i; B_k < F_k, i < k < m-1\} \\
&\quad \dots P\{B_{i+1} < F_{i+1} | B_i \geq F_i\} P\{B_i \geq F_i\} \\
&= (1 - P_{m-1|i}) \dots (1 - P_{i+1|i}) P_i.
\end{aligned}$$

Thus, we can calculate  $P_m$  by unconditioning to yield (20).  $\square$

### Appendix B. The Release-Based Policy.

Let  $B_m$  to be the number of faults when development has been continued without coordination for  $m$  releases. Denote  $C(m, k)$  as the coordination cost at the  $k$ th release when the last coordination is at the  $m$ th release. Then  $C(m, k) = K + bk + cB_{k-m}$ . Also define an array  $u[m]$  to be the optimal cost from the first coordination after the  $m$ th release to the end of the project and an array  $t[m]$  to be the index of the next coordination if we coordinate at the  $m$ th release. The algorithm proceeds as follows:

- 1 Initialize:  $u[M] = 0$ .
- 2 For  $m = M$  and  $m \geq 0$ ,
  - 2.1 Set  $u[m] = C(m, m+1) + u[m+1]$  and  $t[m] = m+1$ .
  - 2.2 For  $k = m+2$  and  $k < M$ , if  $C(m, k) + u[k] < u[m]$ , we set  $u[m] = C(m, k) + u[k]$  and  $t[m] = k$ .
  - 2.3  $m = m+1$ .

At the end of the algorithm,  $u[0]$  returns the total coordination cost.