The utility of the smart jammer at time slot k denoted by  $u_J^{(k)}$  is given by

$$u_J^{(k)} = -\max\left(\frac{Ph_1^{(k)}}{\sigma + y^{(k)}h_3^{(k)}}, \min\left(\frac{Ph_2^{(k)}}{\sigma + y^{(k)}h_4^{(k)}}, \frac{x^{(k)}h_5^{(k)}}{\sigma}\right)\right) + x^{(k)}C_U - y^{(k)}C_J, \tag{1}$$

and the time index k is omitted in the superscript.

**Theorem 1.** The UAV applying Algorithm 1 in the dynamic UAV relay game can achieve the optimal policy (0,0), if

$$\max\left(\frac{Ph_1h_3}{\sigma^2}, \frac{Ph_2h_4}{\sigma^2}, \frac{Ph_2h_4(h_3\sqrt{h_2h_4} - h_4\sqrt{h_1h_3})^2}{\sigma^2(\sqrt{h_1h_3}(1 - h_4) + \sqrt{h_2h_4}(h_3 - 1))^2}\right)$$

$$< C_J \le \frac{Ph_1h_3(h_3\sqrt{h_2h_4} - h_4\sqrt{h_1h_3})^2}{\sigma^2(\sqrt{h_1h_3}(1 - h_4) + \sqrt{h_2h_4}(h_3 - 1))^2}$$
(2)

$$C_U \ge \frac{h_5}{\sigma},$$
 (3)

and the corresponding BER with QPSK is given by

$$P_e = \frac{1}{2} \operatorname{erfc}\left(\frac{Ph_1}{\sigma}\right). \tag{4}$$

*Proof:* By (1), if (3) holds, we have

$$\frac{\partial u(x,0)}{x} = \frac{h_5}{\sigma} - C_U \le 0,\tag{5}$$

thus the utility u is decreases with respect to (w.r.t.) x, and we have

$$u(0,0) = \frac{Ph_1}{\sigma} \ge \max\left(\frac{Ph_1}{\sigma}, \min(\frac{Ph_2}{\sigma}, \frac{xh_5}{\sigma})\right) - xC_U = u(x,0).$$
 (6)

By (1), if (2) holds, we have

$$\frac{\partial u_J(0,y)}{y} = -\max\left(\frac{-Ph_1h_3}{(\sigma + yh_3)^2}, \frac{-Ph_2h_4}{(\sigma + yh_4)^2}\right) - C_J \tag{7}$$

$$\frac{\partial^2 u_J(0,y)}{\partial y^2} = -\max\left(\frac{2Ph_1h_3^2}{(\sigma + yh_3)^3}, \frac{2Ph_2h_4^2}{(\sigma + yh_4)^3}\right) \le 0,\tag{8}$$

indicating that  $u_J(0,y)$  is concave with respect to (w.r.t.) y. Thus,  $\forall y \in [0, P_J^M]$ , we have

$$u_J(0,0) = -\frac{Ph_1}{\sigma} \ge -\frac{Ph_1}{\sigma + yh_3} = u_J(0,y). \tag{9}$$

Thus, the optimal policy is (0,0). According the QPSK modulation, we have

$$P_e = \frac{1}{2} \operatorname{erfc}\left(\frac{Ph_1}{\sigma}\right). \tag{10}$$