The utility of the UAV at time slot k denoted by $u^{(k)}$ is given by

$$u^{(k)} = -\frac{1}{2}\text{erfc}\left(\max\left(\sqrt{\frac{Ph_1^{(k)}}{\sigma + y^{(k)}h_3^{(k)}}}, \min\left(\sqrt{\frac{Ph_2^{(k)}}{\sigma + y^{(k)}h_4^{(k)}}}, \sqrt{\frac{x^{(k)}h_5^{(k)}}{\sigma}}\right)\right)\right) - x^{(k)}C_U, \tag{1}$$

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

$$u_J^{(k)} = \frac{1}{2}\text{erfc}\left(\max\left(\sqrt{\frac{Ph_1^{(k)}}{\sigma + y^{(k)}h_3^{(k)}}}, \min\left(\sqrt{\frac{Ph_2^{(k)}}{\sigma + y^{(k)}h_4^{(k)}}}, \sqrt{\frac{x^{(k)}h_5^{(k)}}{\sigma}}\right)\right)\right) + x^{(k)}C_U - y^{(k)}C_J, \tag{2}$$

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*

$$C_J \geq \frac{3\exp(\frac{-3}{2})h_3}{2\sqrt{\pi}Ph_1} \tag{3}$$

$$C_U \geq \frac{2h_5}{\sqrt{\pi}\sigma}, \tag{4}$$

and the corresponding BER with QPSK is given by

$$P_e = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{Ph_1}{\sigma}}\right). \tag{5}$$

Proof: By (1), if (4) holds, we have

$$\frac{\partial u(x,0)}{x} = \frac{2h_5\exp\left(\frac{-xh_5}{\sigma}\right)}{\sqrt{\pi}\sigma} - C_U, \tag{6}$$

$$\frac{\partial^2 u(x,0)}{\partial x^2} = -\frac{2h_5^2\exp\left(\frac{-xh_5}{\sigma}\right)}{\sqrt{\pi}\sigma^2} \leq 0, \tag{7}$$

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

$$u(0,0) = -\frac{1}{2}\text{erfc}\left(\sqrt{\frac{Ph_1}{\sigma}}\right) \geq -\frac{1}{2}\text{erfc}\left(\max\left(\sqrt{\frac{Ph_1}{\sigma}}, \min\left(\sqrt{\frac{Ph_2}{\sigma}}, \sqrt{\frac{xh_5}{\sigma}}\right)\right)\right) - xC_U = u(x,0). \tag{8}$$

By (2), if (3) holds, we have

$$\frac{\partial u_J(0,y)}{y} = \frac{\sqrt{Ph_1}h_3\exp\left(\frac{-Ph_1}{\sigma+yh_3}\right)}{\sqrt{\pi}(\sigma+yh_3)^{\frac{3}{2}}} - C_J, \tag{9}$$

$$\frac{\partial^2 u_J(0,y)}{\partial y^2} = \frac{\sqrt{Ph_1}h_3^2\left(\frac{Ph_1}{\sigma+yh_3} - \frac{3}{2}\right)\exp\left(\frac{-Ph_1}{\sigma+yh_3}\right)}{\sqrt{\pi}(\sigma+yh_3)^{\frac{5}{2}}} \leq 0, \tag{10}$$

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

$$u_J(0,0) = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{Ph_1}{\sigma}}\right) \geq \frac{1}{2}\text{erfc}\left(\sqrt{\frac{Ph_1}{\sigma+yh_3}}\right) - yC_J = u_J(0,y). \tag{11}$$

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

$$P_e = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{Ph_1}{\sigma}}\right). \tag{12}$$

■