We consider a wireless sensor network (WSN), consisting of several sensors and a fusion center (FC), which is tasked with solving an M-ary hypothesis testing problem. Sensors make M-ary decisions and transmit their digitally modulated decisions over orthogonal channels, which are subject to Rayleigh fading and noise, to the FC. Adopting Bayesian optimality criterion, we consider training and non-training based distributed detection systems and investigate the effect of imperfect channel state information (CSI) on the optimal maximum a posteriori probability (MAP) fusion rules and detection performance, when the sum of training and data symbol transmit powers is fixed. Our results show that, when sensors employ M-ary FSK modulation, the error probability is minimized when training symbol transmit power is zero (regardless of the reception mode at the FC). However, for coherent reception and M-ary PSK modulation the error probability is minimized when half of transmit power is allocated for training symbol.