where $\varepsilon$ is used to tune the operating point of a single observation. In practice, $\varepsilon$ is initialized as $\varepsilon = \frac{1}{15} \sum_{i=1}^{15} s_i + \Delta$, where the first item denotes the current SNR level, and $\Delta$ is a user defined constant. The constant “15” can be set to other value too.

Until now, we can obtain a new feature vector $I_n = \{s_{n-1}, \ldots, s_{n+m}\}^T$ (or $I_n = \{c_{n-1}, \ldots, c_{n+m}\}^T$) from the soft (or hard) decision. Many classifiers based on the new feature could be designed, such as the most simplest calculating the average value of the feature [29], the global hypothesis of the multiple observation [18], the long term amplitude envelope method [22], and the discriminative (weight) training method of the feature [20], [21]. For simplicity, we just calculate the average value of the feature.

$$\Lambda_n = \begin{cases} \frac{1}{1+m+1} \sum_{i=n-1}^{n+m} s_i, & \text{if soft decision is used} \\ \frac{1}{1+m+1} \sum_{i=n-1}^{n+m} c_i, & \text{otherwise} \end{cases}$$

and classify the current observation $o_n$ by

$$o_n \begin{cases} \text{is classified as speech, if } \Lambda_n \geq \eta \\ \text{is classified as noise, otherwise} \end{cases}$$

where $\eta$ is used to tune the operating point of the MO-LLR algorithm.

Fig. 1 shows an example of the detection process of the MO-LLR with $l = m - 1$. From the figure, we could know that when the window length becomes large, the proposed algorithm has a good ability of controlling the randomness of the speech signals but a relatively weak ability of detecting very short pauses between speeches. Therefore, setting the window to a proper length is important to balance the performance between the speech detection accuracy and the false alarm rate. In our realization of the VAD, the hard decision method (6) is adopted, and two thresholds $\eta_{begin}$ and $\eta_{end}$ are used for the beginning-point detection and the ending-point detection of the speeches respectively in stead of a single $\eta$ in (8).

**Algorithm 1: Combining Energy Detection & MO-LLR.**

1: initialization. start from silence.
2: beginning-point (BP) detection:
   if a possible BP $\hat{o}_B$ is detected by Part1 of the energy detection
   if $\hat{o}_B$ is confirmed to be speech by MO-LLR
   search in a range of $(\hat{o}_E - \delta, \hat{o}_E + \delta)$ for the accurate $o_B$ BP by MO-LLR. $o_B$ is defined as the change point from noise to speech.
   goto the ending-point detection (Step 12)
3: else
4: goto the beginning-point detection (Step 18)
5: end
6: else
7: move to next observation, goto Step 2
8: end
9: else
10: move to next observation, goto Step 2
11: end
end

ending-point (EP) detection:
12: if a possible EP $\hat{o}_E$ is detected by Part2 of the energy detection
13: if $\hat{o}_E$ is confirmed to be noise by MO-LLR
14: search in a range of $(\hat{o}_E - \delta, \hat{o}_E + \delta)$ for the accurate EP $o_E$ by MO-LLR. $o_E$ is defined as the change point from speech to noise.
15: if the length from $o_B$ to $o_E$ is too small to be practical
16: delete the detected speech endpoints $o_B$ and $o_E$
17: end
18: goto the BP detection (Step 2)
19: else
20: move to next observation, goto Step 12.
21: end
22: else
23: move to next observation, goto Step 12.
24: end

D. Combination of the Energy Detection Algorithm and the MO-LLR Algorithm

The main consideration of the combination is to detect the noise/speech signals that can be easily differentiated at first, leaving the signals around the endpoints to the MO-LLR sub-algorithm.

Fig. 2 gives a direct explanation of the combination method of the two sub algorithms. From the figure, it’s clear that we just need the MO-LLR sub-algorithm to detect the observations around the possible endpoints that are detected by the energy detection algorithm. Hence, a lot of computation can be saved.

We summarize the proposed algorithm in Algorithm 1 with its state transition graph drawn in Fig. 3. Note that for the MO-LLR sub-algorithm, because an observation might appear not only in the current window but also in the next window when