**Responses to Reviewer 2:**

*Thank you for reviewing our manuscript! Here are our responses to your questions:*

**Question 2:** In the paper, the LLM-based navigation algorithm outperformed the traditional moth-inspired method and the simple rule-based vision and olfaction “Fusion” navigation algorithm in terms of average search time and success rate. Nevertheless, the underlying rationale behind this advantage requires further discussion. For the source localization scenarios set in this study, how do the LLMs' strong multi-modal semantic understanding and reasoning capabilities specifically provide an advantage over the rule-based vision and olfaction “Fusion” navigation algorithm?

***Author’s response***: *Thank you for your feedback!*

*It should be mentioned that our proposed method is a multi-modal navigation algorithm, which means not only it uses olfactory sensors but also visual sensors. Therefore, compared to single-modal navigation algorithms such as the moth-inspired navigation algorithm, our navigation algorithm takes advantage of visual inputs: when the robot sees the odor source visually, it can approach the odor source directly. This visual detection significantly increases the success rate of locating odor sources in both unidirectional and non-unidirectional flow environments. On In addition, the moth-inspired navigation algorithm relies heavily on wind measurements. For instance, the unidirectional flow environment, the moth-inspired method achieves a high success rate (10/16 in Table 2), but in a non-unidirectional flow environment the success rate drops significantly (0/16 in Table 3). In contrast, with the help of visual detection, our method succeeds in both environments.*

*Compared to the rule-based vision and olfaction fusion algorithm (Fusion in Tables 2 and 3), our method utilizes the reasoning and semantic understanding capabilities of LLMs. This skill helps the navigation agent make more intelligent decisions instead of following a set of predefined rules. For instance, when the LLM sees an electrical fan (Query 3 in Fig 11 (c)), it can deduce that the possible odor source location is near the fan’s location. This capability is not achievable for a rule-based Fusion algorithm, which is trained only to recognize visible odor plumes.*

*To clarify this question, we add few sentences in lines 562-571 and 576-582 of section 6.*

**Question 3**: While the authors conducted a total of 64 trials across each search environment, my primary concern is that each method was tested only four times per scenario, which is significantly fewer than the commonly accepted range of 10–15 trials. It appears the authors may not have fully understood this point and have not provided an adequate response.

***Author’s response***: *We appreciate your suggestion!*

*In our experiments, our focus is to investigate how the unidirectional and non-unidirectional flow environments affect the navigation performance of each navigation algorithm. The starting point (the ‘per scenario’ in the question) is not our main focus. It is true that for each starting point, we only tested four trials for each method, but for one environment (unidirectional or non-unidirectional), each method was tested 16 times, which exceeds the commonly acceptable range of 10-15 trials. In our statistical analysis (lines 515-523), we consider factors of airflow environments and navigation methods, not the robot starting points.*

*To emphasize this point, we added lines 410-416 of section 4.2.*

**Question 5**:  "Laminar airflow" and "turbulent airflow" are established technical terms and should not be redefined arbitrarily. I recommend using "unidirectional flow" and "non-unidirectional flow" instead for greater accuracy in terminology.

***Author’s response***: *Thank you for your suggestion!*

*We replaced laminar airflow -> unidirectional flow, and turbulent airflow -> non-unidirectional flow throughout the manuscript. We have labeled the changed places in red color.*