Supplementary materials of the article Response threshold distributions to improve best-of-n decisions in minimalistic robot swarms

Swadhin Agrawal   Sujit P. Baliyarasimhuni
Andreagiovanni Reina

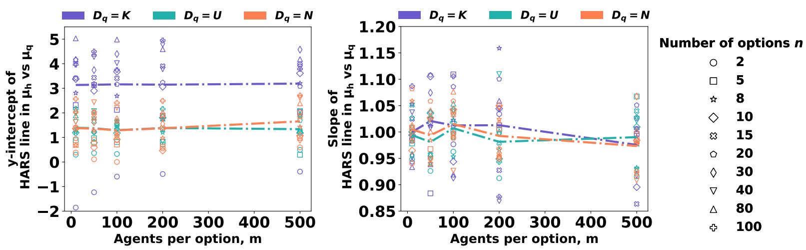

Figure S1: We show that the slope and the y-intercept of the HARS line in μ_h vs μ_q are invariant to changing the number of agents for all the cases of varying number of options in the various types of environment, when the standard deviation of the response thresholds' distribution and the options' quality distribution were fixed at $\sigma_h = \sigma_q = 1$.

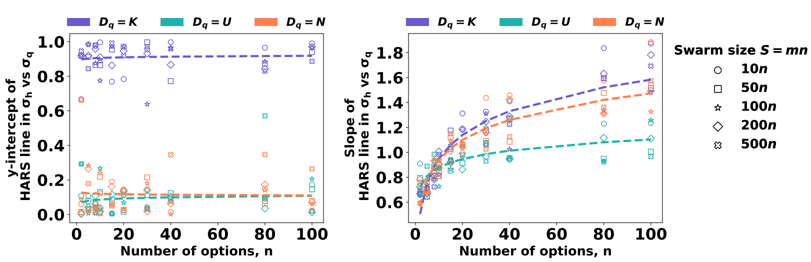

Figure S2: We show that the slope of the HARS line in σ_h vs σ_q varies significantly whereas the y-intercept remains unaffected by changing the number of options in the different type of environments. We run these simulations by naively choosing the response threshold distribution with the same mean as that of the options' quality distribution.

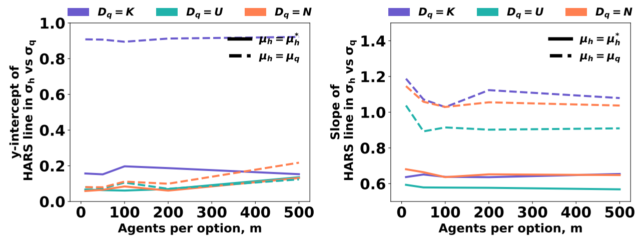

Figure S3: We show the dependency of the slope and the y-intercept of the HARS lines on σ_h vs σ_q plots. We compare the two conditions when $\mu_h = \mu_q$ (dashed lines) and $\mu_h = \mu_h^*$ (solid lines), to estimate how does the slope and the y-intercept get affected by changing the number of agents assigned to each option. We found that both the slope and the y-intercept of the remain invariant to changing the number of agents, m. And this remains valid for all the three types of environments that are under consideration in this study.

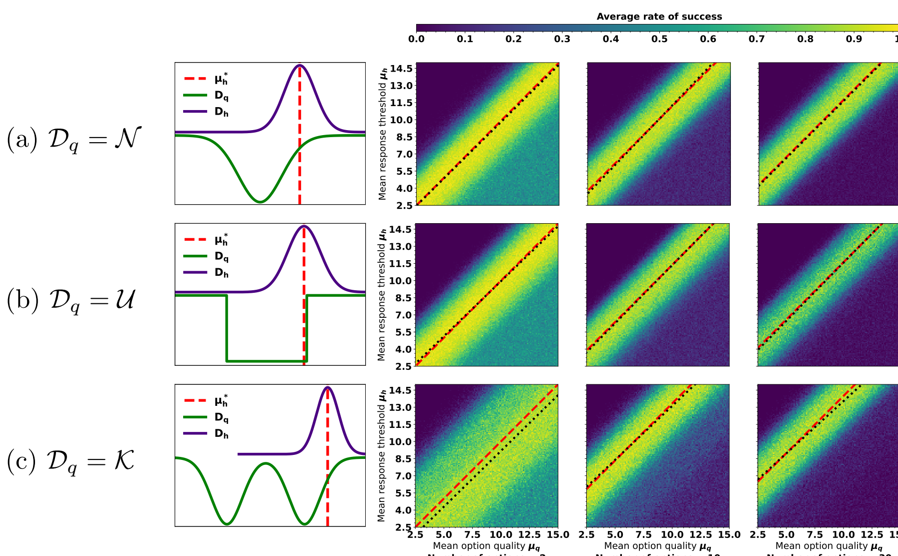

Figure S4: We tested three alternative distributions of options' qualities $\mathcal{D}_q$: (a) Gaussian $\mathcal{N}(\mu_q, \sigma_q)$, (b) uniform $\mathcal{U}(\mu_q, \sigma_q)$, and (c) bimodal distribution $\mathcal{K}(\mu_q, \sigma_q)$. We test the combination of μ_h vs μ_q in the value range $[2.5, 15]$ with a step size of 0.1, for distributions with equal fix standard deviations $\sigma_h = \sigma_q = 1$ and a swarm size of $S = 100n$. The colourmaps show the average rate of success (500 runs per conditions), the dotted black line is the fitted HARS line, and the dashed red line is the best threshold mean μ_q^* predicted with Eq. (1). The two lines match well for all tested types of quality distributions and number of options.

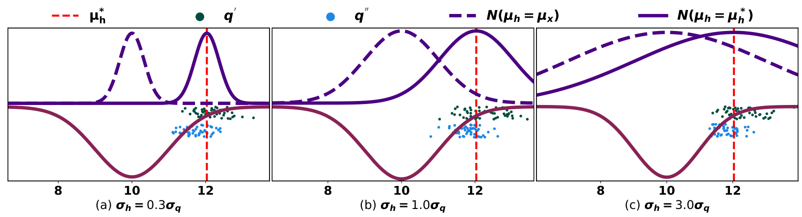

Figure S5: Comparison of the $\mathcal{D}_h = \mathcal{N}$ (top half) and $\mathcal{D}_q = \mathcal{N}$ (bottom half) for increasing ratio of σ_h to σ_q when $\mu_h = \mu_q$ (dashed distribution) and when $\mu_h = \mu_h^*$ (solid distribution). The green and blue dots show the best (q') and second best (q'') value out of 50 sampled points in 50 trials. The red dashed line is μ_h^* computed using Eq. (1) in the main text.

1 Finding the Best Std. dev. σ_h^* for Response Thresholds

Understanding the optimal standard deviation of the response threshold distribution is as important as the mean of the response thresholds' distribution, because the right amount of variability can help minimize the number of robots required to deal with the stochastic nature of the best-of-n decision making process.

Figure S5 shows a schematic overlapping of the options' quality distribution with that of the response thresholds' distribution. The bottom half of the plots shows the options' qualities probability distribution $\mathcal{D}_q = \mathcal{N}(\mu_q, \sigma_q)$, and we report the highest values, q', as a cloud of green points and the second highest value, q'', as blue points, for 50 simulations, assuming the number of options is $n = 50$. The top half of the plots shows the response threshold distribution $\mathcal{D}_h = \mathcal{N}(\mu_h, \sigma_h)$. In each subplot, we show $\mathcal{D}_h$ for two conditions: having the same mean of the options' qualities, $\mu_h = \mu_q$ (dashed distribution), and having the best mean $\mu_h = \mu_h^*$ (solid distribution) computed with Eq. (1).

In the case of $\mu_h = \mu_q$, we find that the standard deviation of response thresholds can play an important role in discerning between high quality options. When the response thresholds' standard deviation is much smaller than the qualities' one, e.g., $\sigma_h = \frac{1}{3}\sigma_q$ in Fig. S5(a), the probability that robots will have their response thresholds between q' and q'' is almost null. Consequently, all high quality options will receive an equal acceptance and be indistinguishable during the voting phase. A standard deviation of the response thresholds higher than the qualities' one, e.g., $\sigma_h = 3\sigma_q$ in Fig. S5(c), can reduce this problem, by spreading more the possible thresholds. While having a large σ_h can improve the system robustness when values of μ_h are suboptimal (e.g., $\mu_h = \mu_q$ for $n > 2$ as in Fig. S5), it is also a waste of resource, because several thresholds are set to values much lower than necessary and only a small percentage of employed robots have a determining role in the collective decision making. Differently, when the mean of the response thresholds is set to a better value, e.g., $\mu_h = \mu_h^*$,

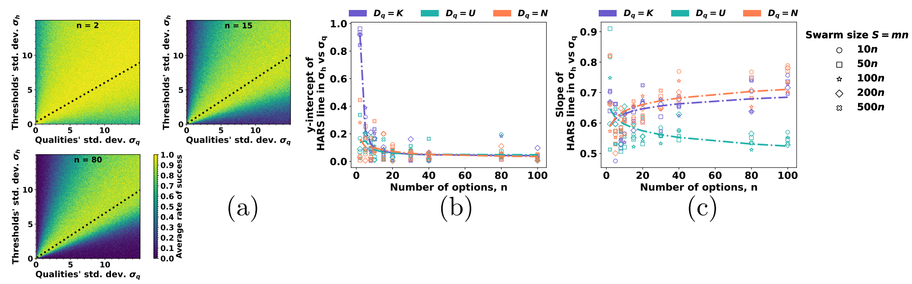

Figure S6: We test the relationship between the standard deviation of the distributions of the option's quality $\mathcal{D}_q = \mathcal{N}(\mu_q = 7.5, \sigma_q)$ and the response thresholds $\mathcal{D}_h = \mathcal{N}(\mu_h = \mu_h^*, \sigma_h)$, i.e., σ_q vs σ_h. The results are the average of 500 simulations per condition in a swarm of $S = 100n$ agents. (a) The colourmaps shows the average rate of success for Gaussian quality distribution. The black dotted HARS lines show that the highest success rate can be obtained with a thresholds' standard deviation smaller than the one of the qualities. Panels (b) and (c) show the y-intercept and slope, respectively, of the fitted HARS line for different swarm sizes and different types of the quality distributions.

the influence of the standard deviation is much reduced. Some variability among response thresholds is always necessary, however the standard deviation can be set to relatively low values, e.g., $\sigma_h = 0.3\sigma_q$ in Fig. S5(a), and still cover the relevant quality range. Here, having unnecessarily high, e.g., $\sigma_h \geq \sigma_q$ in Figs. S5(b-c), does not bring any benefit but only reduces the number of determining robots (i.e., robots with thresholds between q' and q''). This is also visible by the fact that the HARS lines of Figs. S6(a-c) always have a slope smaller than 1, for $\mu_h = \mu_h^*$.

The results of Fig. S6(b-c) numerically confirms the intuition that we just described; they show intercept and slope of the fitted HARS lines for different types of quality distributions and swarm sizes. Differently from Fig. 3, where the slope was approximately constant and the y-intercept largely varied, for σ_h vs σ_q the intercept has negligible values close to zero (except for small $n \leq 3$), while the slope varies. We find mathematical equations, through standard polynomial fitting techniques (NumPy polyfit function [1]), to describe the change in the slopes, with a different equation for each of the considered quality distributions:

$$\sigma_h^* = \begin{cases} (-0.07 \log_{10} n + 0.67)\sigma_q & \text{if } \mathcal{D}_q = \mathcal{U}(\mu_q, \sigma_q) \\ (0.07 \log_{10} n + 0.57)\sigma_q & \text{if } \mathcal{D}_q = \mathcal{N}(\mu_q, \sigma_q) \\ (0.05 \log_{10} n + 0.58)\sigma_q & \text{if } \mathcal{D}_q = \mathcal{K}(\mu_q, \sigma_q)\ . \end{cases} \tag{S1}$$

Because the intercept of the HARS line in σ_h vs σ_q is close to zero (Fig.S6(a)), predicting the slope of the HARS lines is sufficient to determine the relationship between σ_h and σ_q as a function of the number of options n.

It is interesting to note that in environments where the options' qualities are bounded to a maximum value (e.g. $\mathcal{D}_q = \mathcal{U}$), the value of σ_h^* decreases with larger n, as best and second-best options have similar values in a increasingly narrow range. Instead, when qualities values are unbounded (e.g. $\mathcal{D}_q = \mathcal{N}, \mathcal{K}$), the probability of having high-quality outliers increases with n and it seems that a higher standard deviation helps in making correct decision in such cases.

References

[1] Charles R. Harris et al. "Array programming with NumPy". In: *Nature* 585.7825 (2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2.