Supplementary Information

1 Preliminary knowledge of SMC (Sliding mode Control)

Because the change on orientation and velocity of UUV will affect the system matrix like $\mathbf{M}(\mathbf{v}_k)$, its model has strong nonlinear characteristics causing some control strategies fail to stablize the system. In order to obtain a well control accuracy with the nonlinear model under the influence of disturbance τ_k^d, we first employ a sliding mode strategy to stabilize the UUV model.

Design the following sliding surface constraining both position deviation and speed deviation:

$$\mathbf{s}_k = \mathbf{c}\mathbf{v}_k + \eta_k, \tag{1}$$

where $\mathbf{s}_k$ is the weight to balance whether to focus more on the deviation on position or speed. With a sign function $\dot{\mathbf{s}} = -\rho sgn(\mathbf{s})$ play as the reaching law of SMC, and formulate its discrete form:

$$\frac{\mathbf{s}_{k+1} - \mathbf{s}_k}{dt} = -\rho sgn(\mathbf{s}_k). \tag{2}$$

Substitute equations 5 and 1 into equation 2 and simplify them, we can get:

$$\begin{aligned}\tau = &\mathbf{C}(\mathbf{v}_k)\mathbf{v}_k + \mathbf{D}(\mathbf{v}_k)\mathbf{v} + \mathbf{g}(\eta_k) \\ &- \mathbf{M}(\mathbf{v}_k)\mathbf{c}^{-1}\mathbf{J}\mathbf{v} - \mathbf{M}\mathbf{c}^{-1}\rho sgn(\mathbf{s}_k).\end{aligned} \tag{3}$$

Thus the sliding model control law for UUV is obtained. Under this control strategy, the sliding mode state will enter the vicinity of sliding mode surface $s(k) = 0$ with width Δ and stay there forever [1]. That is, $|s(k)| \leq \Delta$ for all $k > k^\star$ with a constant $k^\star$. To get the specific Δ, bring 3 and 2 into the model of UUV and simplify them:

$$\mathbf{s}_{k+1} = \mathbf{s}_k - \rho sgn(\mathbf{s}_k)dt + \mathbf{c}\mathbf{M}(\mathbf{v}_k)^{-1}\tau_k^d dt. \tag{4}$$

As $\mathbf{s}_k$ will eventually approach zero, the bias between $\mathbf{s}_k$ and $\mathbf{s}_{k+1}$ comes from disturbance τ_k^d and the τ to compensate the previous disturbance τ_{k+1}^d. Then it can be concluded that $\{\mathbf{s}_k : |\mathbf{s}_k| \leq \Delta, \Delta = \rho sgn(\mathbf{s}_k) + \mathbf{c}\mathbf{M}^{-1}\tau_k^d dt\}$

2 Preliminary knowledge of LQR (Linear Quartic Optimal Control)

The former strategy, SMC, tends to provide high-gain input with high-frequency oscillations, which is known as "chattering". This can cause excessive wear and tear on actuators and degrade the performance of the control system while increasing the control consumption.

To reach the balance between control accuracy and consumption, a H_2 controller is designed to carry out the task. But since it works for linear system, we need to first linelize the UUV model. Since the control frequency is fast and the UUV's orientation changed slowly, we can view the whole system as a slow time-varying system. Thus, linearize the model of UUV at each point k using Taylor expansion, we can get the linearied model as:

$$\begin{aligned}
\mathbf{v}_{k+1} &= [-\mathbf{M}^{-1}(\mathbf{C}\mathbf{v}_k + \mathbf{D}\mathbf{v}_k + \mathbf{g}) + \mathbf{M}^{-1}(\tau_k + \tau_k^d)]dt + \mathbf{v}_k, \\
\eta_{k+1} &= \mathbf{J}\mathbf{v}_k dt + \eta_k,
\end{aligned} \tag{5}$$

where $\mathbf{M} = \frac{\partial \mathbf{M}(\mathbf{v}_k)}{\partial \mathbf{v}_k}$, $\mathbf{C} = \frac{\partial \mathbf{C}(\mathbf{v}_k)}{\partial \mathbf{v}_k}$, $\mathbf{D} = \frac{\partial \mathbf{D}(\mathbf{v}_k)}{\partial \mathbf{v}_k}$, $\mathbf{g} = \frac{\partial \mathbf{g}(\eta_k)}{\partial \eta_k}$ and $\mathbf{J} = \frac{\partial \mathbf{J}(\eta_k)}{\partial \eta_k}$.

Combine the orientation and velocity into a new vector $\mathbf{x}_{k+1} = [\eta_{k+1}, \mathbf{v}_{k+1}]^T$, we can build the discrete time model for UUV as:

$$\begin{aligned}
\mathbf{x}_{k+1} &= \begin{bmatrix} \mathbf{I}dt & \mathbf{J}dt \\ 0 & -\mathbf{M}^{-1}(\mathbf{C}+\mathbf{D})dt + \mathbf{I}dt \end{bmatrix} \begin{bmatrix} \eta_k \\ \mathbf{v}_k \end{bmatrix} \\
&+ \begin{bmatrix} 0 \\ \mathbf{M}^{-1}dt \end{bmatrix} (\tau_k + \tau_k^d) \\
&= \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k(\tau_k + \tau_k^d).
\end{aligned} \tag{6}$$

where $\mathbf{A}_k \in \Re^{M_1 \times M_1}$ is the system matrix, $\mathbf{B}_k \in \Re^{M_2 \times M_1}$ is the input matrix.

Define the cost function as $F = {\mathbf{x}_k}^{\mathrm{T}} \mathbf{Q} \mathbf{x}_k + {\tau_k}^{\mathrm{T}} \mathbf{R} \tau_k$, and solve a QP(Convex Quadratic Programming) problem using the Riccati equation [2]. We can get the feedback control law as:

$$\tau_k{}^{\star} = -\mathbf{K}_k^o \mathbf{x}_k, \tag{7}$$

$$\mathbf{K}_k^o = -(\mathbf{R} + \mathbf{B}_k^T \mathbf{P}_{k+1}^o \mathbf{B}_k)^{-1} \mathbf{B}_k^T \mathbf{P}_{k+1}^o \mathbf{A}_k, \tag{8}$$

where $\mathbf{K}_k^o$ is the optimal control gain and $\mathbf{P}_k^o$ satisfies a discrete algebra Riccati equation (ARE)

$$\mathbf{P}_k^o = \mathbf{A}_k^T \mathbf{P}_k^o \mathbf{A}_k - \mathbf{A}_k^T \mathbf{P}_k^o \mathbf{B}_k (\mathbf{R} + \mathbf{B}_k^T \mathbf{P}_k^o \mathbf{B}_k)^{-1} \mathbf{B}_k^T \mathbf{P}_k^o \mathbf{A}_k + \mathbf{Q} \tag{9}$$

2.1 Discussion about cooperation Mechanism

In actual marine environment, the UUV system will be subject to different kinds of disturbance. According to their statistical characteristics, they can be divided into two parts. One is the noise that satisfies the Gaussian distribution, which mainly comes from noise formed by the superposition of a large number of independent, small-scale sound sources (such as bubble collapse and raindrop impacts). This kind of disturbance are special because of their fixed mean and variance, leading them can be handled through LQR according to the separation principle [3].

The other can be included as uncertainty disturbance, for example those effects caused by ocean currents and underwater turbulence. The latter kind of disturbance includes most of the interference received by UUV, we unify them only because they do not conform to the Gaussian distribution, such that they can not be dealt with by LQR. In this article, we mainly take three sources of disturbance into consideration, one is those caused by ocean currents, the other is those caused by linearized residual and finally is the model mismatch caused by changes in UUV payload.

To demonstrate the effect caused by Gaussian noise and uncertainty disturbance, we give the mismatched discrete linearlized UUV model:

$$\mathbf{x}_{k+1}^{co} = (\mathbf{A}_k + \Delta\mathbf{A}_k)\mathbf{x}_k^{co} + (\mathbf{B}_k + \Delta\mathbf{B}_k)\tau_k^{co} + \tau_k^d + \mathbf{w}_k, \tag{10}$$

where $\Delta\mathbf{A}_k$ and $\Delta\mathbf{B}_k$ are model mismatch caused by changes in UUV payload, τ_k^d is the ocean current disturbance and $\mathbf{w}_k$ represent Gaussian noise. The above model is the real system model considered in this paper, the other terms remain the same as in model 6.

From the previous introduction, it can be found that LQR can handle the deviation caused by Gaussian disturbance $\mathbf{w}_k$ and SMC can handle the deviation caused by the rest of the disturbance. But when the two sources of deviation mix together, if only apply a single controller (SMC or LQR) to model 10, then there always part of the deviation that can not be properly handled.

To be specific, when SMC is used to stabilize the system, there is often a situation where the deviation has been significantly reduced but still oscillates randomly within a small range due to the presence of Gaussian noise. However, due to the high-frequency and large-input switching characteristics of the SMC, the same drastic control input as the large deviation is applied to the smaller deviation, resulting in overshoot and oscillation. On the other hand, LQR can be viewed as a feedback control strategy with a accurate control gain $\mathbf{K}_k$ designed in the situation of having fully knowledge of the actual model. If there is uncertainty disturbance, LQR will adopt the feedback control strategy with the same control gain. This will result in incomplete compensation, leading to the existence of static error, and even causing the static error to gradually increase until it diverges.

Thus, we need to find a method to withdraw the advantages and eliminate the defects of the above two controllers. The natural idea is to average the output of the two controllers, which is known as mixed controller. However, directly putting these two control inputs together does not mean combining the advantage or overcoming the defects. As is depicted in Fig 1, the left side shows the control inputs of each controller work on themself, causing the LQR to have an improper process on deviation origins from uncertainty disturbance, as well as the SMC to have an improper process on

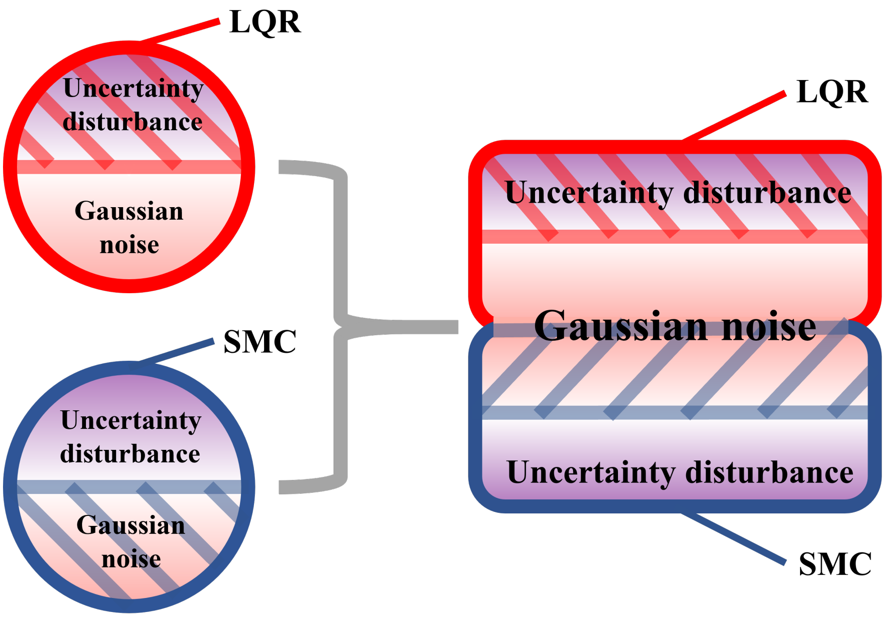

Fig. 1 Why simpliy combine the control input of SMC and LQR do not work

deviation origins from Gaussian noise. However, with both improper control inputs kept, their results on the right side led to no deviation being dealt with properly.

Naturally, unlike the previous method of directly combining the two controllers, the idea of decomposing the deviation into two parts suitable for LQR and SMC comes to our mind. Make an identity transformation on 10 formulates:

$$\begin{aligned}\mathbf{x}_{k+1}^{co} &= (\mathbf{A}_k + \Delta\mathbf{A}_k)\mathbf{x}_k^r + (\mathbf{B}_k + \Delta\mathbf{B}_k)\tau_k^r + \tau_k^d \\ &+ \mathbf{A}_k\mathbf{x}_k^o + \mathbf{B}_k\tau_k^o + \mathbf{w}_k \\ &= \mathbf{x}_{k+1}^r + \mathbf{x}_{k+1}^o, \end{aligned} \tag{11}$$

where $\mathbf{x}_k^{co}$, $\mathbf{x}_k^r$ and $\mathbf{x}_k^o$ are both state vectors representing the posture deviation with the relationship as $\mathbf{x}_k^{co} = \mathbf{x}_k^o + \mathbf{x}_k^r$. As is displayed in model 11, the posrue deviation is divided into two parts. The first only include the uncertainty disturbance while the other only contains the Gaussian noise. As shown in Fig 2, through a deviation seperation algorithm, the mixed deviation can be divided into two independent deviation according to its origins. Then, we can use LQR to solve the part coming from Gaussian noise and SMC to solve the part coming from uncertainty disturbance. Compared to the mixed controller, the controller we proposed can truly complement each other's strengths, so we call it COC (collaborative controller).

From Fig 3, the basic framework of control policy is still feedback control and the feedback item we can use is still the posture deviation. Mathematically speaking, a deviation separation mechanism initially partitions the deviation into distinct parts

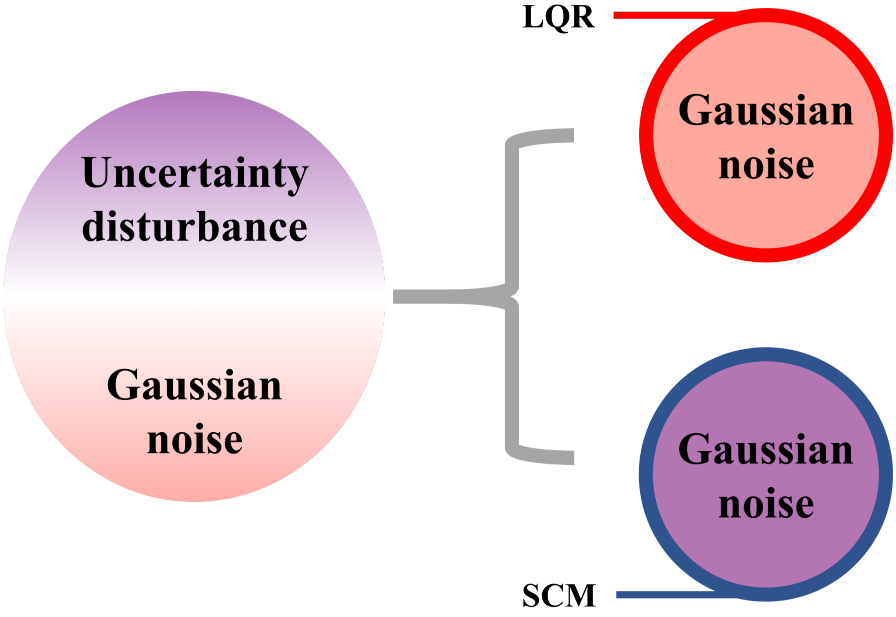

Fig. 2 Why collaboration stragety works

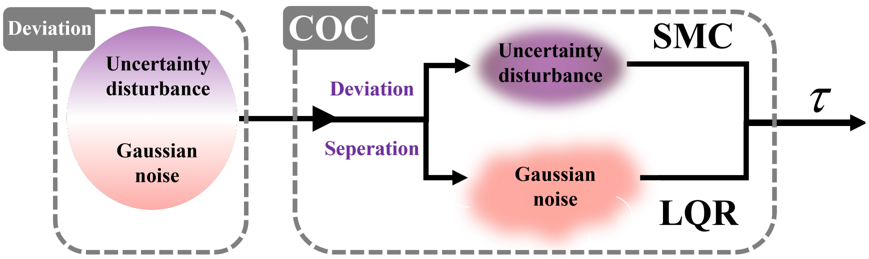

Fig. 3 The framework of COC

according to their statistical attributes. Subsequently, LQR addresses the term from Gaussian noise, while SMC manages the other terms from uncertainty disturbance.

References

[1] Ma, H., Wu, J., Xiong, Z.: Discrete-time sliding-mode control with improved quasi-sliding-mode domain. IEEE Transactions on Industrial Electronics **63**(10), 6292–6304 (2016)

[2] Tijjani, A.S., Chemori, A., Creuze, V.: A survey on tracking control of unmanned underwater vehicles: Experiments-based approach. Annual Reviews in Control **54**, 125–147 (2022)

[3] Anderson, B.D.O., Moore, J.B.: Optimal Control: Linear Quadratic Methods. Prentice Hall, Englewood Cliffs, NJ (1990)

[4] Levine, S.: Reinforcement learning and control as probabilistic inference: Tutorial and review. arXiv preprint arXiv:1805.00909 (2018)

[5] Cook, R.D.: Assessment of local influence. Journal of the Royal Statistical Society Series B: Statistical Methodology **48**(2), 133–155 (1986)

[6] Yu, S., Xue, W., Luan, X., Liu, F.: A real-time performance evaluation indicators of lqg controller for systems with unknown changes. IEEE Transactions on Instrumentation and Measurement **73**, 1–12 (2024) https://doi.org/10.1109/TIM.2024.3421434