Scalable Multi-Robot Active Exploration

George P. Kontoudis
Maryland Robotics Center, University of Maryland
Email: kont@umd.edu

I. Introduction

Exploration with a team of robots is an important task for the majority of robotic missions, including environmental monitoring [1], search and rescue [2], and coverage path planning [3]. An efficient way to perform robotic exploration is by using active learning [4] with Gaussian processes (GPs) [5]–[7]. Although GPs have proven to be accurate in various multi-robot applications [8]–[16], they suffer from scalability issues. In addition, for their implementation in decentralized networks, the communication overhead becomes significantly high. Since coordination of multi-robot systems requires effective communication to maintain network connectivity, the communication channel cannot be congested for the implementation of GPs. Typically, communication links are considered as deterministic distances between robots [17]–[26]. Unambiguously, the communication performance depends on the distance of robots, yet the environment imposes significant challenges. The environmental variations lead to complex channel modeling which cannot be reliably represented by deterministic functions, but rather are interpreted as a statistical realization [27]. Yet, even the best statistical model of communication performance cannot incorporate every single source of noise or all parameters of the environment, and may fail in practice. Probabilistic learning combined with model-based elements are useful to accurately model the communication performance. Thus, GPs are not only powerful to explore the environment, but also can create communication maps for the interaction between robots. After the goal locations are specified to efficiently explore the environment, a motion planning technique is required. Since the obstacle space is usually unknown *a priori*, a rapid replanning algorithm is needed to adjust in the obstacle space changes [28]. In addition, kinodynamic constraints have to be satisfied with a low-level controller. In such cases, an optimal control scheme is desired, yet it requires significant computations [29], suffering from uncertain dynamics and external disturbances [30].

Objective: We aim to develop scalable and decentralized GPs for the exploration of non-stationary environments using a team of robots. To maintain network connectivity, we intend to predict the communication performance at unvisited sites. For the navigation, we seek real-time optimal kinodynamic motion planning techniques that are robust to external disturbances.

Related Work: To alleviate the computation demand for GP training, a factorized method is discussed in [31]. In [32], the alternating direction method of multipliers (ADMM) [33] is used to solve the distributed optimization problem of GP training. To further reduce the complexity, the inexact proximal ADMM [34] is employed in [35]. However, all these works require a centralized topology to achieve hyper-parameter GP training. Scalable approaches for GP prediction involve local approximations [36]. The goal is to aggregate local predictions to a central server [31], [37]–[39]. Since most of these methods are inconsistent [31], [37], [38], a nested pointwise aggregation of experts is introduced in [40], but with high computations in the central node. A generalized approach that equips local datasets with a global dataset is proposed in [41]. Similarly to the training methods, local approximations for GP prediction impose a centralized topology, and thus are impractical for multi-robot missions [18].

A method to generate global and local underwater acoustic communication maps based partially on kriging (equivalent to GPs [6]) is proposed in [42]. In [43], communication maps of known terrestrial environments are built using GPs from multiple agents. Optimal relay positioning for improving indoor communication of mobile sensor networks using GPs is discussed in [44]. Prediction is achieved using model-free ordinary kriging that assumes a stationary random field.

Kinodynamic motion planning in static environments with optimal control for controllable linear systems is introduced in [45]. A feedback kinodynamic motion planning technique for static environments that combines convex optimization tools and Lyapunov functions for linearized systems is discussed in [46]. In [47], a variant of deep reinforcement learning is combined with RRT for kinodynamic motion planning. In [48], the authors propose a real-time kinodynamic motion planner using local replanning and offline machine learning techniques to facilitate an online implementation. All of these methods require significant offline computations, exact knowledge of the system dynamics, and they are not robust to external disturbances.

II. Technical Approach and Contributions

The proposed method for scalable and fully decentralized exploration with a team of robots consists three steps: i) decentralized GPs for the variable of interest and communication performance prediction; ii) decentralized active learning with GPs for the variable of interest; and iii) kinodynamic motion planning with continuous-time Q-learning for low-level control of each robot, as shown in Fig. 1.

A. Decentralized Gaussian Processes [49], [50]

1) Training: Factorized GP training [31] invokes independence between local models $\mathcal{M}_i$ to result in the approxima-

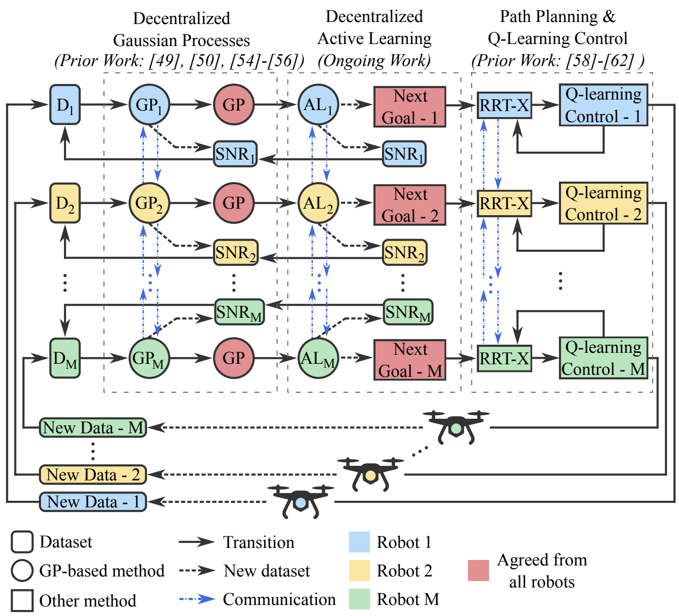

Fig. 1. Structure of the scalable multi-robot active exploration with GPs.

tion of the log-likelihood function as $\mathcal{L}(\boldsymbol{\theta}) \approx \sum_{i=1}^{M} \mathcal{L}_i(\boldsymbol{\theta}_i)$, where M is the number of robots. This also implies that the costly inversion of the covariance matrix is approximated as $\boldsymbol{K}^{-1} \approx \mathrm{diag}(\boldsymbol{K}_1^{-1}, \ldots, \boldsymbol{K}_M^{-1})$, reducing the computations from $\mathcal{O}(N^3)$ to $\mathcal{O}(N^3/M^3)$, where $N = \sum_{i=1}^{M} N_i$ is the number of total observations. Assuming that the decentralized network topology is strongly connected, the maximum likelihood optimization problem for the estimation of GP hyper-parameters $\boldsymbol{\theta}$ is expressed as the edge ADMM formulation [51]. To address this problem we consider a first-order approximation of the local log-likelihood around the optimizing hyper-parameter vector at the s-th iteration $\boldsymbol{\theta}^{(s)}$ as discussed in [52]. We extend the latter approach by computing an analytical solution of the nested ADMM optimization problem. The results reveal three orders of magnitude faster GP training when compared to a typical GP [6]. In addition, it is reported similar time execution and accuracy with the centralized methods [31], [35], [41] for all fleet sizes.

2) Prediction: The estimated hyper-parameter vector $\hat{\boldsymbol{\theta}}$ is employed from each robot i to compute a local mean $\mu_i(\boldsymbol{x}_*)$ and variance $\sigma_i^2(\boldsymbol{x}_*)$. Then, the robots exchange local values to perform local approximations. We provide decentralized algorithms for the implementation of the centralized local approximations [31], [37]–[39], [41] using consensus protocols [53]. Essentially, each robot i employs the average consensus values and the size of the fleet to replace all summations in the local approximations. Next, we introduce a methodology to select statistically correlated nearest neighbors for the location of interest $\boldsymbol{x}_*$ by formulating a random process over the local mean values, which is showed to be a GP $(\mu_1, \ldots, \mu_M, y(\boldsymbol{x}_*))^\intercal \sim \mathcal{GP}(\mu_\mu, \boldsymbol{C}_{\theta,\mu})$. We leverage the cross-covariance elements of $\boldsymbol{C}_{\theta,\mu}$ to represent the correlation of each robot i to the location of interest. The results demonstrate convergence of all decentralized methods to the centralized local approximations, yielding identical predictions. Moreover, the nearest neighbor method indicates 42.5% reduction in robot involvement to the local approximations without sacrificing prediction accuracy.

B. Communication Performance Prediction [54]–[56]

The underwater acoustic (UWA) communication of submarine robots is considered as a non-stationary, spatial random field. We use signal-to-noise ratio (SNR) to evaluate the communication performance. The SNR measurement model is separated to a deterministic part for the mean function and a Gaussian random field for the spatial variability. Inspired by the UWA stochastic model [27], we design a basis function to estimate the mean function, which incorporates a linear-log relationship to the distance of robots r as $\boldsymbol{X} = [1, \Phi(x), \Phi(y), r, \log r]$, where $\Phi(\cdot)$ is a radial basis function. For the spatial variability we employ three covariance functions and select the most suitable model by using the posterior distribution of the Bayesian information criterion (BIC) [57]. When no clear preference to a model is indicated, a new covariance function is constructed as the convex combination of the three models, with weights the posterior of BIC. Field trials illustrate accurate predictions and realistic uncertainty quantification in various ambient noise environments.

C. Online Kinodynamic Motion Planning [58]–[62]

A scalable decoupled approach is designed, where: i) the sampling-based path planning method RRT$^{\mathrm{X}}$ [28], [63] provides a set of boundary value problems in dynamic environments; and ii) a Q-learning controller is responsible for low-level robot navigation. The continuous-time Q-learning controller is formulated to solve the finite-horizon optimal control problem with completely unknown system dynamics. The Q-function is defined as the optimal value function and the Hamiltonian corresponding to the finite-horizon objective $\mathcal{Q}(x; u; t) \coloneqq V^\star(x; t) + \mathcal{H}(x; u; \partial V^\star/\partial t, \partial V^\star/\partial x)$, as in [64]. We form an actor-critic structure on the error of the Q-function and the error of the optimal control [65]. Then, we learn the actor-critic weights by using gradient descent with closed-form derivatives. Lyapunov-based proofs ensure closed-loop stability of the equilibrium point. The results reveal collision-free navigation in unknown dynamic obstacle environments.

III. Future Work

In the future, we plan to decentralize the implementation of active learning (AL) with GPs [4], [5]. The idea of AL is to explore an environment with sequential GP model updates. Objective functions are used as criteria to identify the next goal sampling locations. Their main disadvantage is the scalability for real-time implementation, especially with non-myopic criteria [66]. In addition, for decentralized AL the robots need to communicate as little as possible. Although fast GP updates can be used to alleviate the computational complexity [67], analytical solutions for the partial derivatives of the objective functions can accelerate the convergence rate and reduce the communication overhead of decentralized methods.

References

[1] K.-C. Ma, L. Liu, H. K. Heidarsson, and G. S. Sukhatme, "Data-driven learning and planning for environmental sampling," *Journal of Field Robotics*, vol. 35, no. 5, pp. 643–661, 2018.

[2] J. P. Queralta, J. Taipalmaa, B. C. Pullinen, V. K. Sarker, T. N. Gia, H. Tenhunen, M. Gabbouj, J. Raitoharju, and T. Westerlund, "Collaborative multi-robot search and rescue: Planning, coordination, perception, and active vision," *IEEE Access*, vol. 8, pp. 191 617–191 643, 2020.

[3] R. Almadhoun, T. Taha, L. Seneviratne, and Y. Zweiri, "A survey on multi-robot coverage path planning for model reconstruction and mapping," *SN Applied Sciences*, vol. 1, no. 8, pp. 1–24, 2019.

[4] S. Seo, M. Wallat, T. Graepel, and K. Obermayer, "Gaussian process regression: Active data selection and test point rejection," in *Mustererkennung 2000*. Springer, 2000, pp. 27–34.

[5] R. B. Gramacy, *Surrogates: Gaussian Process Modeling, Design and Optimization for the Applied Sciences*. Boca Raton, Florida: Chapman Hall/CRC, 2020, http://bobby.gramacy.com/surrogates/.

[6] C. E. Rasmussen and C. K. Williams, *Gaussian Processes for Machine Learning*, 2nd ed. Cambridge, MA, USA: MIT Press, 2006.

[7] N. Cressie, *Statistics for spatial data*, 2nd ed. Wiley, 1993.

[8] Y. Xu, J. Choi, and S. Oh, "Mobile sensor network navigation using Gaussian processes with truncated observations," *IEEE Transactions on Robotics*, vol. 27, no. 6, pp. 1118–1131, 2011.

[9] A. Singh, A. Krause, C. Guestrin, and W. J. Kaiser, "Efficient informative sensing using multiple robots," *Journal of Artificial Intelligence Research*, vol. 34, pp. 707–755, 2009.

[10] G. Pillonetto, L. Schenato, and D. Varagnolo, "Distributed multi-agent Gaussian regression via finite-dimensional approximations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 9, pp. 2098–2111, 2018.

[11] T. N. Hoang, Q. M. Hoang, K. H. Low, and J. How, "Collective online learning of Gaussian processes in massive multi-agent systems," in *AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, 2019, pp. 7850–7857.

[12] M. Tavassolipour, S. A. Motahari, and M. T. M. Shalmani, "Learning of Gaussian processes in distributed and communication limited systems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 8, pp. 1928–1941, 2020.

[13] D. Jang, J. Yoo, C. Y. Son, D. Kim, and H. J. Kim, "Multi-robot active sensing and environmental model learning with distributed Gaussian process," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5905–5912, 2020.

[14] M. E. Kepler and D. J. Stilwell, "An approach to reduce communication for multi-agent mapping applications," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2020, pp. 4814–4820.

[15] Z. Yuan and M. Zhu, "Communication-aware distributed Gaussian process regression algorithms for real-time machine learning," in *American Control Conference*, 2020, pp. 2197–2202.

[16] V. Suryan and P. Tokekar, "Learning a spatial field in minimum time with a team of robots," *IEEE Transactions on Robotics*, vol. 36, no. 5, pp. 1562–1576, 2020.

[17] W. Burgard, M. Moors, C. Stachniss, and F. E. Schneider, "Coordinated multi-robot exploration," *IEEE Transactions on Robotics*, vol. 21, no. 3, pp. 376–386, 2005.

[18] F. Bullo, J. Cortes, and S. Martinez, *Distributed control of robotic networks: A mathematical approach to motion coordination algorithms*. Princeton University Press, 2009, vol. 27.

[19] J. Cortés, "Distributed Kriged Kalman filter for spatial estimation," *IEEE Transactions on Automatic Control*, vol. 54, no. 12, pp. 2816–2827, 2009.

[20] R. K. Williams and G. S. Sukhatme, "Constrained interaction and coordination in proximity-limited multiagent systems," *IEEE Transactions on Robotics*, vol. 29, no. 4, pp. 930–944, 2013.

[21] L. V. Nguyen, S. Kodagoda, R. Ranasinghe, and G. Dissanayake, "Information-driven adaptive sampling strategy for mobile robotic wireless sensor network," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 1, pp. 372–379, 2015.

[22] Y. Kantaros and M. M. Zavlanos, "Distributed intermittent connectivity control of mobile robot networks," *IEEE Transactions on Automatic Control*, vol. 62, no. 7, pp. 3109–3121, 2016.

[23] R. Khodayi-mehr, Y. Kantaros, and M. M. Zavlanos, "Distributed state estimation using intermittently connected robot networks," *IEEE Transactions on Robotics*, vol. 35, no. 3, pp. 709–724, 2019.

[24] M. Minelli, J. Panerati, M. Kaufmann, C. Ghedini, G. Beltrame, and L. Sabattini, "Self-optimization of resilient topologies for fallible multi-robots," *Robotics and Autonomous Systems*, vol. 124, p. 103384, 2020.

[25] X. Yu and M. A. Hsieh, "Synthesis of a time-varying communication network by robot teams with information propagation guarantees," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1413–1420, 2020.

[26] Y. Sung, A. K. Budhiraja, R. K. Williams, and P. Tokekar, "Distributed assignment with limited communication for multi-robot multi-target tracking," *Autonomous Robots*, vol. 44, no. 1, pp. 57–73, 2020.

[27] P. Qarabaqi and M. Stojanovic, "Statistical characterization and computationally efficient modeling of a class of underwater acoustic communication channels," *IEEE Journal of Oceanic Engineering*, vol. 38, no. 4, pp. 701–717, 2013.

[28] M. Otte and E. Frazzoli, "RRT^X: Asymptotically optimal single-query sampling-based motion planning with quick replanning," *The International Journal of Robotics Research*, vol. 35, no. 7, pp. 797–822, 2016.

[29] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal Control*, 3rd ed. John Wiley & Sons, 2012.

[30] F. Berkenkamp and A. P. Schoellig, "Safe and robust learning control with Gaussian processes," in *European Control Conference*, 2015, pp. 2496–2501.

[31] M. Deisenroth and J. W. Ng, "Distributed Gaussian processes," in *International Conference on Machine Learning*, 2015, pp. 1481–1490.

[32] Y. Xu, F. Yin, W. Xu, J. Lin, and S. Cui, "Wireless traffic prediction with scalable Gaussian process: Framework, algorithms, and verification," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 6, pp. 1291–1306, 2019.

[33] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers*, 2011, vol. 3, no. 1.

[34] M. Hong, Z.-Q. Luo, and M. Razaviyayn, "Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems," *SIAM Journal on Optimization*, vol. 26, no. 1, pp. 337–364, 2016.

[35] A. Xie, F. Yin, Y. Xu, B. Ai, T. Chen, and S. Cui, "Distributed Gaussian processes hyperparameter optimization for big data using proximal ADMM," *IEEE Signal Processing Letters*, vol. 26, no. 8, pp. 1197–1201, 2019.

[36] H. Liu, Y.-S. Ong, X. Shen, and J. Cai, "When Gaussian process meets big data: A review of scalable GPs," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 11, pp. 4405–4423, 2020.

[37] G. E. Hinton, "Training products of experts by minimizing contrastive divergence," *Neural computation*, vol. 14, no. 8, pp. 1771–1800, 2002.

[38] V. Tresp, "A Bayesian committee machine," *Neural computation*, vol. 12, no. 11, pp. 2719–2741, 2000.

[39] Y. Cao and D. J. Fleet, "Generalized product of experts for automatic and principled fusion of Gaussian process predictions," *arXiv preprint arXiv:1410.7827*, 2014.

[40] D. Rullière, N. Durrande, F. Bachoc, and C. Chevalier, "Nested Kriging predictions for datasets with a large number of observations," *Statistics and Computing*, vol. 28, no. 4, pp. 849–867, 2018.

[41] H. Liu, J. Cai, Y. Wang, and Y. S. Ong, "Generalized robust Bayesian committee machine for large-scale Gaussian process regression," in *International Conference on Machine Learning*, 2018, pp. 3131–3140.

[42] D. Horner and G. Xie, "Data-driven acoustic communication modeling for undersea collaborative navigation," in *IEEE OCEANS-San Diego*, 2013, pp. 1–10.

[43] A. Q. Li, P. K. Penumarthi, J. Banfi, N. Basilico, J. M. O'Kane, I. Rekleitis, S. Nelakuditi, and F. Amigoni, "Multi-robot online sensing strategies for the construction of communication maps," *Autonomous Robots*, vol. 44, pp. 299–319, 2019.

[44] J. Kim, P. Ladosz, and H. Oh, "Optimal communication relay positioning in mobile multi-node networks," *Robotics and Autonomous Systems*, vol. 129, p. 103517, 2020.

[45] D. J. Webb and J. van den Berg, "Kinodynamic RRT*: Asymptotically optimal motion planning for robots with linear dynamics," in *IEEE Int. Conf. on Robotics and Automation*, 2013, pp. 5054–5061.

[46] R. Tedrake, I. R. Manchester, M. Tobenkin, and J. W. Roberts, "LQR-Trees: Feedback motion planning via sums-of-squares verification," *The International Journal of Robotics Research*, vol. 29, no. 8, pp. 1038–1052, 2010.

[47] H.-T. L. Chiang, J. Hsu, M. Fiser, L. Tapia, and A. Faust, “RL-RRT: Kinodynamic motion planning via learning reachability estimators from RL policies,” *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 4298–4305, 2019.
[48] R. E. Allen and M. Pavone, “A real-time framework for kinodynamic planning in dynamic environments with application to quadrotor obstacle avoidance,” *Robotics and Autonomous Systems*, vol. 115, pp. 174–193, 2019.
[49] G. P. Kontoudis and D. J. Stilwell, “Decentralized nested Gaussian processes for multi-robot systems,” in *IEEE International Conference on Robotics and Automation*, 2021, pp. 8881–8887.
[50] G. P. Kontoudis and D. Stilwell, “Fully decentralized, scalable Gaussian processes for multi-agent federated learning,” *arXiv preprint arXiv:2203.02865*, 2022.
[51] W. Shi, Q. Ling, K. Yuan, G. Wu, and W. Yin, “On the linear convergence of the ADMM in decentralized consensus optimization,” *IEEE Transactions on Signal Processing*, vol. 62, no. 7, pp. 1750–1761, 2014.
[52] T.-H. Chang, M. Hong, and X. Wang, “Multi-agent distributed optimization via inexact consensus ADMM,” *IEEE Transactions on Signal Processing*, vol. 63, no. 2, pp. 482–497, 2014.
[53] R. Olfati-Saber, J. A. Fax, and R. M. Murray, “Consensus and cooperation in networked multi-agent systems,” *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.
[54] G. P. Kontoudis and D. J. Stilwell, “A comparison of kriging and cokriging for estimation of underwater acoustic communication performance,” in *International Conference on Underwater Networks & Systems*, 2019, pp. 1–8.
[55] G. P. Kontoudis, S. Krauss, and D. J. Stilwell, “Model-based learning of underwater acoustic communication performance for marine robots,” *Robotics and Autonomous Systems*, vol. 142, p. 103811, 2021.
[56] G. P. Kontoudis and D. J. Stilwell, “Prediction of acoustic communication performance in marine robots using model-based kriging,” in *American Control Conference*, 2021, pp. 3779–3786.
[57] A. A. Neath and J. E. Cavanaugh, “The Bayesian information criterion: Background, derivation, and applications,” *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 4, no. 2, pp. 199–203, 2012.
[58] G. P. Kontoudis and K. G. Vamvoudakis, “Robust kinodynamic motion planning using model-free game-theoretic learning,” in *American Control Conference*, 2019, pp. 273–278.
[59] G. P. Kontoudis, Z. Xu, and K. G. Vamvoudakis, “Online, model-free motion planning in dynamic environments: An intermittent, finite horizon approach with continuous-time Q-learning,” in *American Control Conference*, 2020, pp. 3873–3878.
[60] G. P. Kontoudis and K. G. Vamvoudakis, “Kinodynamic motion planning with continuous-time Q-learning: An online, model-free, and safe navigation framework,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 12, pp. 3803–3817, 2019.
[61] J. Netter, G. P. Kontoudis, and K. G. Vamvoudakis, “Bounded rational RRT-QX: Multi-agent motion planning in dynamic human-like environments using cognitive hierarchy and Q-learning,” in *IEEE Conference on Decision and Control*, 2021, pp. 3597–3602.
[62] G. P. Kontoudis, K. G. Vamvoudakis, and Z. Xu, “RRT-QX: Real-time kinodynamic motion planning in dynamic environments with continuous-time reinforcement learning,” in *Brain and Cognitive Intelligence: Control in Robotics*, B. Wei, Ed. Taylor & Francis Group, CRC Press.
[63] M. Otte and E. Frazzoli, “RRT^{X}: Real-time motion planning/replanning for environments with unpredictable obstacles,” in *Algorithmic Foundations of Robotics XI*, 2015, pp. 461–478.
[64] K. G. Vamvoudakis, “Q-learning for continuous-time linear systems: A model-free infinite horizon optimal control approach,” *Systems & Control Letters*, vol. 100, pp. 14–20, 2017.
[65] B. Kiumarsi, K. G. Vamvoudakis, H. Modares, and F. L. Lewis, “Optimal and autonomous control using reinforcement learning: A survey,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 6, pp. 2042–2062, 2018.
[66] D. A. Cohn, Z. Ghahramani, and M. I. Jordan, “Active learning with statistical models,” *Journal of Artificial Intelligence Research*, vol. 4, pp. 129–145, 1996.
[67] R. B. Gramacy and H. K. Lee, “Adaptive design and analysis of supercomputer experiments,” *Technometrics*, vol. 51, no. 2, pp. 130–145, 2009.