

References

- F. Archetti, A. Candelieri *Bayesian Optimization and Data Science* (Springer, Berlin, 2019)
- R. Astudillo, P. Frazier, Bayesian optimization of composite functions, in *International Conference on Machine Learning* (PMLR, New York, 2019), pp. 354–363
- M. Balandat, B. Karrer, D. Jiang, S. Daulton, B. Letham, A.G. Wilson, E. Bakshy, Botorch: a framework for efficient monte-carlo Bayesian optimization. *Adv. Neural Inf. Proces. Syst.* **33**, 21524–21538 (2020)
- R. Bansal, B. Samanta, S. Dalmia, N. Gupta, S. Vashishth, S. Ganapathy, A. Bapna, P. Jain, P. Talukdar, LLM augmented LLMS: expanding capabilities through composition. *arXiv preprint arXiv:2401.02412* (2024)
- M. Baratchi, C. Wang, S. Limmer, J.N. van Rijn, H. Hoos, T. Bäck, M. Olhofer, Automated machine learning: past, present and future. *Artif. Intell. Rev.* **57**(5), 122 (2024)
- R. Bardenet, M. Brendel, B. Kégl, M. Sebag, Collaborative hyperparameter tuning, in *International Conference on Machine Learning* (PMLR, New York, 2013), pp. 199–207
- S. Barocas, M. Hardt, A. Narayanan, *Fairness in Machine Learning* (NIPS Tutorial, California, 2017)
- S. Belakaria, A. Deshwal, Max-value entropy search for multi-objective bayesian optimization, in *International Conference on Neural Information Processing Systems (NeurIPS)* (2019)
- S. Belakaria, A. Deshwal, J.R. Doppa, Multi-fidelity multi-objective bayesian optimization: an output space entropy search approach, in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34 (2020a), pp. 10035–10043
- S. Belakaria, A. Deshwal, N.K. Jayakodi, J.R. Doppa, Uncertainty-aware search framework for multi-objective bayesian optimization, in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34 (2020b), pp. 10044–10052
- F. Berkenkamp, A.P. Schoellig, A. Krause, No-regret bayesian optimization with unknown hyperparameters. *J. Mach. Learn. Res.* **20**(50), 1–24 (2019)
- J. Blank, K. Deb, Pymoo: multi-objective optimization in python. *IEEE Access* **8**, 89497–89509 (2020)
- I. Bogunovic, A. Krause, Misspecified gaussian process bandit optimization. *Adv. Neural Inf. Proces. Syst.* **34**, 3004–3015 (2021)
- E.V. Bonilla, K. Chai, C. Williams, Multi-task gaussian process prediction. *Adv. Neural Inf. Proces. Syst.* **20** (2007)
- D.E. Bryson, M.P. Rumpfkeil, Variable-fidelity surrogate modeling of lambda wing transonic aerodynamic performance, in *54th AIAA Aerospace Sciences Meeting* (2016), p. 0294
- A.D. Bull, Convergence rates of efficient global optimization algorithms. *J. Mach. Learn. Res.* **12**(10) (2011)

- J. Buolamwini, T. Gebru, Gender shades: intersectional accuracy disparities in commercial gender classification, in *Conference on Fairness, Accountability and Transparency* (PMLR, New York, 2018), pp. 77–91
- A. Candelieri, F. Archetti, Miso-wildcosts: multi information source optimization with location dependent costs. *arXiv preprint arXiv:2102.04951* (2021a)
- A. Candelieri, F. Archetti, Sparsifying to optimize over multiple information sources: an augmented gaussian process based algorithm. *Struct. Multidiscip. Optim.* **64**, 239–255 (2021b)
- A. Candelieri, D. Soldi, F. Archetti, Cost-effective sensors placement and leak localization—the Neptun pilot of the ice water project. *J. Water Supply Res Technol.—AQUA* **64**(5), 567–582 (2015)
- A. Candelieri, R. Perego, I. Giordani, A. Ponti, F. Archetti, Modelling human active search in optimizing black-box functions. *Soft Comput.* **24**(23), 17771–17785 (2020)
- A. Candelieri, R. Perego, F. Archetti, Green machine learning via augmented gaussian processes and multi-information source optimization. *Soft Comput.* **25**(19), 12591–12603 (2021)
- A. Candelieri, A. Ponti, F. Archetti, Explaining exploration–exploitation in humans. *Big Data and Cognitive Computing* **6**(4), 155 (2022)
- A. Candelieri, A. Ponti, F. Archetti, Uncertainty quantification and exploration–exploitation trade-off in humans. *J. Ambient. Intell. Humaniz. Comput.*, 1–34 (2023a)
- A. Candelieri, A. Ponti, F. Archetti, Wasserstein enabled bayesian optimization of composite functions. *J. Ambient. Intell. Humaniz. Comput.* **14**(8), 11263–11271 (2023b)
- A. Candelieri, A. Ponti, F. Archetti, Fair and green hyperparameter optimization via multi-objective and multiple information source bayesian optimization. *Mach. Learn.* **113**(5), 2701–2731 (2024)
- Y. Chai, S. Wang, Y. Sun, H. Tian, H. Wu, H. Wang, Clip-tuning: towards derivative-free prompt learning with a mixture of rewards. *arXiv preprint arXiv:2210.12050* (2022)
- R. Charayron, T. Lefebvre, N. Bartoli, J. Morlier. Towards a multi-fidelity & multi-objective bayesian optimization efficient algorithm. *Aerospace Science and technology*, 142: 108673, 2023.
- A. Chaudhuri, A.N. Marques, R. Lam, K.E. Willcox, Reusing information for multifidelity active learning in reliability-based design optimization, in *AIAA Scitech 2019 Forum* (2019), p. 1222
- L. Chen, J. Chen, T. Goldstein, H. Huang, T. Zhou, Instructzero: efficient instruction optimization for black-box large language models. *arXiv preprint arXiv:2306.03082* (2023)
- Y. Chen, X. Song, C. Lee, Z. Wang, R. Zhang, D. Dohan, K. Kawakami, G. Kochanski, A. Doucet, M. Ranzato, et al., Towards learning universal hyperparameter optimizers with transformers. *Adv. Neural Inf. Proces. Syst.* **35**, 32053–32068 (2022)
- S.R. Chowdhury, A. Gopalan, No-regret algorithms for multi-task bayesian optimization, in *International Conference on Artificial Intelligence and Statistics* (PMLR, New York, 2021), pp. 1873–1881
- A.F. Cruz, M. Hardt, Unprocessing seven years of algorithmic fairness. *arXiv preprint arXiv:2306.07261* (2023)
- L. Csató, M. Opper, Sparse representation for gaussian process models, in *Advances in Neural Information Processing Systems* (2001), pp. 444–450
- L. Csató, M. Opper, Sparse on-line gaussian processes. *Neural Comput.* **14**(3), 641–668 (2002)
- J. Cuesta Ramirez, R. Le Riche, O. Roustant, G. Perrin, C. Durantin, A. Gliere, A comparison of mixed-variables bayesian optimization approaches. *Adv. Model. Simul. Eng. Sci.* **9**(1), 6 (2022)
- S. Dai, J. Song, Y. Yue, Multi-task bayesian optimization via gaussian process upper confidence bound, in *ICML 2020 Workshop on Real World Experiment Design and Active Learning* (2020), pp. 1–12
- C. Deng, H. Ding, J. Miao, F. Tan, Co-kriging model-based multi-fidelity regression for refining milling stability predictions of variable tool overhang length using limited experimental data. *Int. J. Prod. Res.*, 1–22 (2024)

- M. Deng, J. Wang, C.-P. Hsieh, Y. Wang, H. Guo, T. Shu, M. Song, E.P. Xing, Z. Hu, Rlprompt: optimizing discrete text prompts with reinforcement learning. *arXiv preprint arXiv:2205.12548* (2022)
- P. Dhar, The carbon impact of artificial intelligence. *Nat. Mach. Intell.* **2**(8), 423–425 (2020)
- S. Diao, Z. Huang, R. Xu, X. Li, Y. Lin, X. Zhou, T. Zhang, Black-box prompt learning for pre-trained language models. *arXiv preprint arXiv:2201.08531* (2022)
- P.A. Dowd, E. Pardo-Igúzquiza, The many forms of co-kriging: A diversity of multivariate spatial estimators. *Math. Geosci.* **56**(2), 387–413 (2024)
- D. Eriksson, M. Pearce, J. Gardner, R.D. Turner, M. Poloczek, Scalable global optimization via local bayesian optimization. *Adv. Neural Inf. Proces. Syst.* **32** (2019)
- P. Feliot, J. Bect, E. Vazquez, A bayesian approach to constrained single-and multi-objective optimization. *J. Glob. Optim.* **67**(1-2), 97–133 (2017)
- M. Feurer, B. Letham, E. Bakshy, Scalable meta-learning for bayesian optimization using ranking-weighted gaussian process ensembles, in *AutoML Workshop at ICML*, vol. 7 (2018), p. 5
- A. Forrester, A. Sobester, A. Keane, *Engineering Design via Surrogate Modelling: A Practical Guide* (Wiley, New York, 2008)
- A.I. Forrester, A. Sobester, A.J. Keane, Multi-fidelity optimization via surrogate modelling. *Proc. R. Soc. A: Math. Phys. Eng. Sci.* **463**(2088), 3251–3269 (2007)
- Z.Z. Foumani, M. Shishehbor, A. Yousefpour, R. Bostanabad, Multi-fidelity cost-aware bayesian optimization. *Comput. Methods Appl. Mech. Eng.* **407**, 115937 (2023)
- P.I. Frazier, Bayesian optimization, in *Recent Advances in Optimization and Modeling of Contemporary Problems* (Informs, New York, 2018), pp. 255–278
- P.I. Frazier, W.B. Powell, S. Dayanik, A knowledge-gradient policy for sequential information collection. *SIAM J. Control. Optim.* **47**(5), 2410–2439 (2008)
- S.A. Friedler, C. Scheidegger, S. Venkatasubramanian, S.C.E.P. Hamilton, D. Roth, A comparative study of fairness-enhancing interventions in machine learning, in *Proceedings of the Conference on Fairness, Accountability, and Transparency* (2019), pp. 329–338
- T. Furuya, M.V. de Hoop, G. Peyré, Transformers are universal in-context learners. *arXiv preprint arXiv:2408.01367* (2024)
- R. Garnett, *Bayesian Optimization* (Cambridge University, Cambridge, 2023)
- E.C. Garrido-Merchán, D. Hernández-Lobato, Dealing with categorical and integer-valued variables in bayesian optimization with gaussian processes. *Neurocomputing* **380**, 20–35 (2020)
- S. Ghoreishi, W. Thomison, D. Allaire, Sequential information-theoretic and reification-based approach for querying multi-information sources. *J. Aerosp. Inf. Syst.* **16**(12), 575–587 (2019a)
- S.F. Ghoreishi, D. Allaire, Multi-information source constrained bayesian optimization. *Struct. Multidiscip. Optim.* **59**, 977–991 (2019)
- S.F. Ghoreishi, A. Molkeri, R. Arróyave, D. Allaire, A. Srivastava, Efficient use of multiple information sources in material design. *Acta Materialia* **180**, 260–271 (2019b)
- O. Giustolisi, A. Simone, L. Ridolfi, Network structure classification and features of water distribution systems. *Water Resour. Res.* **53**(4), 3407–3423 (2017)
- R.B. Gramacy, *Surrogates: Gaussian Process Modeling, Design, and Optimization for the Applied Sciences* (Chapman and Hall/CRC, New York, 2020)
- Q. Guo, R. Wang, J. Guo, B. Li, K. Song, X. Tan, G. Liu, J. Bian, Y. Yang, Connecting large language models with evolutionary algorithms yields powerful prompt optimizers. *arXiv preprint arXiv:2309.08532* (2023)
- H.C. Herbol, M. Poloczek, P. Clancy, Cost-effective materials discovery: Bayesian optimization across multiple information sources. *Mater. Horiz.* **7**(8), 2113–2123 (2020)
- T. Hofmann, B. Schölkopf, A.J. Smola, *Kernel Methods in Machine Learning* (2008)
- M. Hort, Z. Chen, J.M. Zhang, F. Sarro, M. Harman, Bias mitigation for machine learning classifiers: A comprehensive survey. *arXiv preprint arXiv:2207.07068* (2022)
- E.J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, W. Chen, Lora: low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685* (2021)
- F. Hutter, L. Kotthoff, J. Vanschoren, *Automated Machine Learning: Methods, Systems, Challenges* (Springer Nature, Berlin, 2019)

- F. Irshad, S. Karsch, A. Döpp, Expected hypervolume improvement for simultaneous multi-objective and multi-fidelity optimization. *arXiv preprint arXiv:2112.13901* (2021)
- K. Jamieson, A. Talwalkar, Non-stochastic best arm identification and hyperparameter optimization, in *Artificial Intelligence and Statistics* (PMLR, New York, 2016), pp. 240–248
- H. Jin, X. Han, J. Yang, Z. Jiang, Z. Liu, C.-Y. Chang, H. Chen, X. Hu, LLM maybe LONGLM: Self-extend LLM context window without tuning. *arXiv preprint arXiv:2401.01325* (2024)
- T.T. Joy, S. Rana, S. Gupta, S. Venkatesh, A flexible transfer learning framework for bayesian optimization with convergence guarantee. *Expert Syst. Appl.* **115**, 656–672 (2019)
- K. Kandasamy, G. Dasarathy, J.B. Oliva, J. Schneider, B. Póczos, Gaussian process bandit optimisation with multi-fidelity evaluations. *Adv. Neural Inf. Proces. Syst.* **29** (2016)
- K. Kandasamy, G. Dasarathy, J. Oliva, J. Schneider, B. Poczos, Multi-fidelity gaussian process bandit optimisation. *J. Artif. Intell. Res.* **66**, 151–196 (2019)
- S. Keerthi, W. Chu, A matching pursuit approach to sparse gaussian process regression. *Adv. Neural Inf. Proces. Syst.* **18** (2005)
- M.C. Kennedy, A. O'Hagan, Predicting the output from a complex computer code when fast approximations are available. *Biometrika* **87**(1), 1–13 (2000)
- D. Khatamsaz, L. Peddareddygar, S. Friedman, D.L. Allaire, Efficient multi-information source multiobjective Bayesian optimization. in *AIAA Scitech 2020 Forum*, (2020)
- D. Khatamsaz, B. Vela, R. Arróyave, Multi-objective bayesian alloy design using multi-task gaussian processes. *Mater. Lett.* **351**, 135067 (2023)
- A. Klein, S. Falkner, S. Bartels, P. Hennig, F. Hutter, Fast bayesian optimization of machine learning hyperparameters on large datasets, in *Artificial Intelligence and Statistics* (PMLR, New York, 2017), pp. 528–536
- K.A. Klise, M. Bynum, D. Moriarty, R. Murray, A software framework for assessing the resilience of drinking water systems to disasters with an example earthquake case study. *Environ. Model Softw.* **95**, 420–431 (2017)
- R. Lam, D.L. Allaire, K.E. Willcox, Multifidelity optimization using statistical surrogate modeling for non-hierarchical information sources, in *56th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference* (2015), p. 0143
- R. Lange, Y. Tian, Y. Tang, Evolution transformer: In-context evolutionary optimization, in *Proceedings of the Genetic and Evolutionary Computation Conference Companion* (2024), pp. 575–578
- T. Le Quy, A. Roy, V. Iosifidis, W. Zhang, E. Ntoutsis, A survey on datasets for fairness-aware machine learning. *Wiley Interdiscip. Rev.: Data Min. Knowl. Discovery* **12**(3), e1452 (2022)
- J. Lehman, J. Gordon, S. Jain, K. Ndousse, C. Yeh, K.O. Stanley, Evolution through large models, in *Handbook of Evolutionary Machine Learning* (Springer, Berlin, 2023), pp. 331–366
- B. Lester, R. Al-Rfou, N. Constant, The power of scale for parameter-efficient prompt tuning. *arXiv preprint arXiv:2104.08691* (2021)
- L. Li, K. Jamieson, G. DeSalvo, A. Rostamizadeh, A. Talwalkar, Hyperband: a novel bandit-based approach to hyperparameter optimization. *J. Mach. Learn. Res.* **18**(1), 6765–6816 (2017)
- P. Li, J. Yang, M.A. Islam, S. Ren, Making AI less “thirsty”: Uncovering and addressing the secret water footprint of ai models (2025). <https://arxiv.org/abs/2304.03271>
- X.L. Li, P. Liang, Prefix-tuning: Optimizing continuous prompts for generation. *arXiv preprint arXiv:2101.00190* (2021)
- Y.L. Li, T.G. Rudner, A.G. Wilson, A study of bayesian neural network surrogates for bayesian optimization. *arXiv preprint arXiv:2305.20028* (2023)
- X. Lin, Z. Wu, Z. Dai, Y. Shu, P. Jaillet, B.K.H. Low, Use your instinct: Instruction optimization using neural bandits coupled with transformers. *arXiv preprint arXiv:2310.02905* (2023)
- S. Liu, C. Chen, X. Qu, K. Tang, Y.-S. Ong, Large language models as evolutionary optimizers, in *2024 IEEE Congress on Evolutionary Computation (CEC)* (IEEE, New York, 2024a), pp. 1–8
- T. Liu, N. Astorga, N. Seedat, M. van der Schaar, Large language models to enhance bayesian optimization. *arXiv preprint arXiv:2402.03921* (2024b)
- G. Luo, D. Han, Y. Zhang, H. Ruan, A digital twin for advancing battery fast charging based on a bayesian optimization-based method. *J. Energy Storage* **93**, 112365 (2024)

- L. Mainini, A. Serani, M.P. Rumpfkeil, E. Minisci, D. Quagliarella, H. Pehlivan, S. Yildiz, S. Ficini, R. Pellegrini, F. Di Fiore, et al., Analytical benchmark problems for multifidelity optimization methods. *arXiv preprint arXiv:2204.07867* (2022)
- A. March, K. Willcox, Provably convergent multifidelity optimization algorithm not requiring high-fidelity derivatives. *AIAA J.* **50**(5), 1079–1089 (2012)
- A. Marco, F. Berkenkamp, P. Hennig, A.P. Schoellig, A. Krause, S. Schaal, S. Trimpe, Virtual vs. real: trading off simulations and physical experiments in reinforcement learning with bayesian optimization, in *2017 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, New York, 2017), pp. 1557–1563
- A. Marques, R. Lam, K. Willcox, Contour location via entropy reduction leveraging multiple information sources. *Adv. Neural Inf. Proces. Syst.* **31** (2018)
- N. Maus, Z.J. Lin, M. Balandat, E. Bakshy, Joint composite latent space bayesian optimization. *arXiv preprint arXiv:2311.02213* (2023)
- N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, A. Galstyan, A survey on bias and fairness in machine learning. *ACM Comput. Surv. (CSUR)* **54**(6), 1–35 (2021)
- P. Mikkola, J. Martinelli, L. Filstroff, S. Kaski, Multi-fidelity bayesian optimization with unreliable information sources, in *International Conference on Artificial Intelligence and Statistics* (PMLR, New York, 2023), pp. 7425–7454
- S. Mirchandani, F. Xia, P. Florence, B. Ichter, D. Driess, M.G. Arenas, K. Rao, D. Sadigh, A. Zeng, Large language models as general pattern machines. *arXiv preprint arXiv:2307.04721* (2023)
- H.B. Moss, D.S. Leslie, J. Gonzalez, P. Rayson, Gibbon: general-purpose information-based bayesian optimisation. *J. Mach. Learn. Res.* **22**(235), 1–49 (2021a)
- H.B. Moss, D.S. Leslie, P. Rayson, Mumbo: multi-task max-value bayesian optimization, in *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2020, Ghent, Belgium, September 14–18, 2020, Proceedings, Part III* (Springer, Berlin, 2021b), pp. 447–462
- W. Neiswanger, L. Yu, S. Zhao, C. Meng, S. Ermon, Generalizing bayesian optimization with decision-theoretic entropies. *Adv. Neural Inf. Proces. Syst.* **35**, 21016–21029 (2022)
- G. Nguyen, S. Biswas, H. Rajan, Fix fairness, don't ruin accuracy: Performance aware fairness repair using automl, in *Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering* (2023), pp. 502–514
- M. Nobar, J. Keller, A. Rupenyan, M. Khosravi, J. Lygeros, Guided bayesian optimization: data-efficient controller tuning with digital twin, in *IEEE Transactions on Automation Science and Engineering* (2024)
- L. Papenmeier, L. Nardi, M. Poloczek, Increasing the scope as you learn: Adaptive bayesian optimization in nested subspaces. *Adv. Neural Inf. Proces. Syst.* **35**, 11586–11601 (2022)
- L. Papenmeier, L. Nardi, M. Poloczek, Bounce: reliable high-dimensional bayesian optimization for combinatorial and mixed spaces. *Adv. Neural Inf. Proces. Syst.* **36**, 1764–1793 (2023)
- B. Paria, K. Kandasamy, B. Póczos, A flexible framework for multi-objective bayesian optimization using random scalarizations, in *Uncertainty in Artificial Intelligence* (PMLR, New York, 2020), pp. 766–776
- M. Pearce, J. Branke, Continuous multi-task bayesian optimisation with correlation. *Eur. J. Oper. Res.* **270**(3), 1074–1085 (2018)
- B. Peherstorfer, B. Kramer, K. Willcox, Combining multiple surrogate models to accelerate failure probability estimation with expensive high-fidelity models. *J. Comput. Phys.* **341**, 61–75 (2017)
- V. Perrone, M. Donini, K. Kenthapadi, C. Archambeau, Bayesian optimization with fairness constraints, in *International Conference on Machine Learning (Automated Machine Learning Workshop)* (2020)
- V. Perrone, M. Donini, M.B. Zafar, R. Schmucker, K. Kenthapadi, C. Archambeau, Fair bayesian optimization, in *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (2021), pp. 854–863
- D. Pessach, E. Shmueli, A review on fairness in machine learning. *ACM Comput. Surv. (CSUR)* **55**(3), 1–44 (2022)

- M. Poloczek, J. Wang, P. Frazier, Multi-information source optimization. *Adv. Neural Inf. Proces. Syst.* **30** (2017)
- A. Ponti, A. Candelieri, F. Archetti, A new evolutionary approach to optimal sensor placement in water distribution networks. *Water* **13**(12), 1625 (2021)
- A. Prasad, P. Hase, X. Zhou, M. Bansal, Grips: gradient-free, edit-based instruction search for prompting large language models. *arXiv preprint arXiv:2203.07281* (2022)
- R. Pryzant, D. Iter, J. Li, Y.T. Lee, C. Zhu, M. Zeng, Automatic prompt optimization with “gradient descent” and beam search. *arXiv preprint arXiv:2305.03495* (2023)
- J. Quinonero-Candela, C.E. Rasmussen, A unifying view of sparse approximate gaussian process regression. *J. Mach. Learn. Res.* **6**, 1939–1959 (2005)
- C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, P.J. Liu, Exploring the limits of transfer learning with a unified text-to-text transformer. *J. Mach. Learn. Res.* **21**(140), 1–67 (2020)
- A. Ramachandran, S. Gupta, S. Rana, S. Venkatesh, Information-theoretic multi-task learning framework for bayesian optimisation, in *AI 2019: Advances in Artificial Intelligence: 32nd Australasian Joint Conference, Adelaide, SA, Australia, December 2–5, 2019, Proceedings 32* (Springer, Berlin, 2019), pp. 497–509
- M.C. Ramos, S.S. Michtavy, M.D. Porosoff, A.D. White, Bayesian optimization of catalysts with in-context learning. *arXiv preprint arXiv:2304.05341* (2023)
- B. Ru, A. Alvi, V. Nguyen, M.A. Osborne, S. Roberts, Bayesian optimisation over multiple continuous and categorical inputs, in *International Conference on Machine Learning* (PMLR, New York, 2020), pp. 8276–8285
- A. Sabbatella, A. Ponti, A. Candelieri, F. Archetti, Bayesian optimization using simulation-based multiple information sources over combinatorial structures. *Mach. Learn. Knowl. Extr.* **6**(4), 2232–2247 (2024a)
- A. Sabbatella, A. Ponti, I. Giordani, A. Candelieri, F. Archetti, Prompt optimization in large language models. *Mathematics* **12**(6), 929 (2024b)
- I. Salehin, M.S. Islam, P. Saha, S. Noman, A. Tuni, M.M. Hasan, M.A. Baten, Automl: a systematic review on automated machine learning with neural architecture search. *J. Inf. Intell.* **2**(1), 52–81 (2024)
- R. Schmucker, M. Donini, V. Perrone, M.B. Zafar, C. Archambeau, Multi-objective multi-fidelity hyperparameter optimization with application to fairness, in *NeurIPS Workshop on Meta-Learning*, vol. 2 (2020)
- J. Schreiter, D. Nguyen-Tuong, M. Toussaint, Efficient sparsification for gaussian process regression. *Neurocomputing* **192**, 29–37 (2016)
- R. Schwartz, J. Dodge, N.A. Smith, O. Etzioni, Green AI. *Commun. ACM* **63**(12), 54–63 (2020)
- M. Scutari, F. Panero, M. Proissl, Achieving fairness with a simple ridge penalty. *arXiv preprint arXiv:2105.13817* (2021)
- M. Seeger, F. Steinke, K. Tsuda, Bayesian inference and optimal design in the sparse linear model, in *Artificial Intelligence and Statistics* (PMLR, New York, 2007), pp. 444–451
- M.W. Seeger, C.K. Williams, N.D. Lawrence, Fast forward selection to speed up sparse gaussian process regression, in *International Workshop on Artificial Intelligence and Statistics* (PMLR, New York, 2003), pp. 254–261
- R. Sen, K. Kandasamy, S. Shakkottai, Multi-fidelity black-box optimization with hierarchical partitions, in *International Conference on Machine Learning* (PMLR, New York, 2018), pp. 4538–4547
- O. Sener, V. Koltun, Multi-task learning as multi-objective optimization. *Adv. Neural Inf. Proces. Syst.* **31** (2018)
- B. Shahriari, K. Swersky, Z. Wang, R.P. Adams, N. De Freitas, Taking the human out of the loop: a review of bayesian optimization. *Proc. IEEE* **104**(1), 148–175 (2015)
- P. Shaw, J. Uszkoreit, A. Vaswani, Self-attention with relative position representations. *arXiv preprint arXiv:1803.02155* (2018)
- J. Shawe-Taylor, Kernel methods for pattern analysis. Cambridge University Press Google Schola **2**, 181–201 (2004)

- A. Shehabi, A. Hubbard, A. Newkirk, N. Lei, M.A.B. Siddik, B. Holecek, J. Koomey, E. Masanet, D. Sartor, et al. 2024 united states data center energy usage report (2024)
- S. Shen, L. Hou, Y. Zhou, N. Du, S. Longpre, J. Wei, H.W. Chung, B. Zoph, W. Fedus, X. Chen, et al., Mixture-of-experts meets instruction tuning: A winning combination for large language models. *arXiv preprint arXiv:2305.14705* (2023)
- W. Shi, X. Han, H. Gonen, A. Holtzman, Y. Tsvetkov, L. Zettlemoyer, Toward human readable prompt tuning: Kubrick's the shining is a good movie, and a good prompt too? *arXiv preprint arXiv:2212.10539* (2022)
- T. Shin, Y. Razeghi, R.L. Logan IV, E. Wallace, S. Singh, Autoprompt: eliciting knowledge from language models with automatically generated prompts. *arXiv preprint arXiv:2010.15980* (2020)
- A.J. Smola, P.L. Bartlett, Sparse greedy gaussian process regression, in *Advances in Neural Information Processing Systems* (2001), pp. 619–625
- J. Song, Y. Chen, Y. Yue, A general framework for multi-fidelity bayesian optimization with gaussian processes, in *The 22nd International Conference on Artificial Intelligence and Statistics* (PMLR, New York, 2019), pp. 3158–3167
- X. Song, O. Li, C. Lee, B. Yang, D. Peng, S. Perel, Y. Chen, Omnipred: language models as universal regressors. *arXiv preprint arXiv:2402.14547* (2024a)
- X. Song, Y. Tian, R.T. Lange, C. Lee, Y. Tang, Y. Chen, Position paper: leveraging foundational models for black-box optimization: benefits, challenges, and future directions. *arXiv preprint arXiv:2405.03547* (2024b)
- N. Srinivas, A. Krause, S.M. Kakade, M.W. Seeger, Information-theoretic regret bounds for gaussian process optimization in the bandit setting. *IEEE Trans. Inf. Theory* **58**(5), 3250–3265 (2012)
- E. Strubell, A. Ganesh, A. McCallum, Energy and policy considerations for modern deep learning research, in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34 (2020), pp. 13693–13696
- J. Su, M. Ahmed, Y. Lu, S. Pan, W. Bo, Y. Liu, Roformer: enhanced transformer with rotary position embedding. *Neurocomputing* **568**, 127063 (2024)
- Q. Sun, T. Chen, S. Liu, J. Chen, H. Yu, B. Yu, Correlated multi-objective multi-fidelity optimization for HLS directives design. *ACM Trans. Des. Autom. Electron. Syst. (TODAES)* **27**(4), 1–27 (2022a)
- T. Sun, Z. He, H. Qian, Y. Zhou, X. Huang, X. Qiu, BBTv2: towards a gradient-free future with large language models. *arXiv preprint arXiv:2205.11200* (2022b)
- T. Sun, Y. Shao, H. Qian, X. Huang, X. Qiu, Black-box tuning for language-model-as-a-service, in *International Conference on Machine Learning* (PMLR, New York, 2022c), pp. 20841–20855
- S. Suzuki, S. Takeno, T. Tamura, K. Shitara, M. Karasuyama, Multi-objective bayesian optimization using pareto-frontier entropy, in *International Conference on Machine Learning* (PMLR, New York, 2020), pp. 9279–9288
- J. Svenson, T. Santner, Multiobjective optimization of expensive-to-evaluate deterministic computer simulator models. *Comput. Stat. Data Anal.* **94**, 250–264 (2016)
- K. Swersky, J. Snoek, R.P. Adams, Multi-task bayesian optimization. *Adv. Neural Inf. Proces. Syst.* **26** (2013)
- S. Takeno, H. Fukuoka, Y. Tsukada, T. Koyama, M. Shiga, I. Takeuchi, M. Karasuyama, Multi-fidelity bayesian optimization with max-value entropy search and its parallelization, in *International Conference on Machine Learning* (PMLR, New York, 2020), pp. 9334–9345
- R.D. Tordecilla, A.A. Juan, J.R. Montoya-Torres, C.L. Quintero-Araujo, J. Panadero, Simulation-optimization methods for designing and assessing resilient supply chain networks under uncertainty scenarios: a review. *Simul. Model. Pract. Theory* **106**, 102166 (2021)
- T. Tornede, A. Tornede, J. Hanselle, F. Mohr, M. Wever, E. Hüllermeier, Towards green automated machine learning: Status quo and future directions. *J. Artif. Intell. Res.* **77**, 427–457 (2023)
- K. Vafa, J. Chen, J. Kleinberg, S. Mullainathan, A. Rambachan, Evaluating the world model implicit in a generative model. *arxiv 2024. arXiv preprint arXiv:2406.03689* (2024)
- A. Vaswani, Attention is all you need. *Advances in Neural Information Processing Systems* (2017)

- R. Verdecchia, J. Sallou, L. Cruz, A systematic review of green AI. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **13**(4), e1507 (2023)
- S. Verma, J. Rubin, Fairness definitions explained, in *2018 IEEE/ACM International Workshop on Software Fairness (Fairware)* (IEEE, New York, 2018), pp. 1–7
- K.P. Wabersich, M. Toussaint, Advancing bayesian optimization: the mixed-global-local (MGL) kernel and length-scale cool down. *arXiv preprint arXiv:1612.03117* (2016)
- J. Wackers, M. Visonneau, S. Ficini, R. Pellegrini, A. Serani, M. Diez, Adaptive n-fidelity metamodels for noisy CFD data, in *AIAA AVIATION 2020 FORUM* (2020), p. 3161
- G. Wahba, *Spline Models for Observational Data*, vol. 59 (SIAM, New York, 1990)
- J. Wang, N. Chiang, A. Gillette, J.L. Peterson, A multifidelity bayesian optimization method for inertial confinement fusion design. *Phys. Plasmas* **31**(3) (2024)
- L. Wang, E. Demeulemeester, Simulation optimization in healthcare resource planning: a literature review. *IIE Trans.* **55**(10), 985–1007 (2023)
- Z. Wang, N. de Freitas, Theoretical analysis of bayesian optimisation with unknown gaussian process hyper-parameters. *arXiv preprint arXiv:1406.7758* (2014)
- Z. Wang, S. Jegelka, Max-value entropy search for efficient bayesian optimization, in *International Conference on Machine Learning* (PMLR, New York, 2017), pp. 3627–3635
- Z. Wang, C. Gehring, P. Kohli, S. Jegelka, Batched large-scale bayesian optimization in high-dimensional spaces. *arXiv preprint arXiv:1706.01445* (2018)
- H. Weerts, F. Pfisterer, M. Feurer, K. Eggensperger, E. Bergman, N. Awad, J. Vanschoren, M. Pechenizkiy, B. Bischl, F. Hutter, Can fairness be automated? guidelines and opportunities for fairness-aware automl. *arXiv preprint arXiv:2303.08485* (2023)
- J. Wei, X. Wang, D. Schuurmans, M. Bosma, F. Xia, E. Chi, Q.V. Le, D. Zhou, et al., Chain-of-thought prompting elicits reasoning in large language models. *Adv. Neural Inf. Proces. Syst.* **35**, 24824–24837 (2022)
- Y. Wen, N. Jain, J. Kirchenbauer, M. Goldblum, J. Geiping, T. Goldstein, Hard prompts made easy: gradient-based discrete optimization for prompt tuning and discovery. *Adv. Neural Inf. Proces. Syst.* **36** (2024)
- L. While, L. Bradstreet, L. Barone, A fast way of calculating exact hypervolumes. *IEEE Trans. Evol. Comput.* **16**(1), 86–95 (2011)
- C.K. Williams, C.E. Rasmussen, *Gaussian Processes for Machine Learning*, vol. 2 (MIT Press, Cambridge, MA, 2006)
- M. Wistuba, J. Grabocka, Few-shot bayesian optimization with deep kernel surrogates. *arXiv preprint arXiv:2101.07667* (2021)
- J. Wu, P. Frazier, The parallel knowledge gradient method for batch bayesian optimization. *Adv. Neural Inf. Proces. Syst.* **29** (2016)
- J. Wu, P.I. Frazier, *Continuous-Fidelity Bayesian Optimization with Knowledge Gradient* (2018)
- J. Wu, S. Toscano-Palmerin, P.I. Frazier, A.G. Wilson, Practical multi-fidelity bayesian optimization for hyperparameter tuning, in *Uncertainty in Artificial Intelligence* (PMLR, New York, 2020), pp. 788–798
- G. Xiao, Y. Tian, B. Chen, S. Han, M. Lewis, Efficient streaming language models with attention sinks. *arXiv preprint arXiv:2309.17453* (2023)
- C. Yang, X. Wang, Y. Lu, H. Liu, Q.V. Le, D. Zhou, X. Chen, Large language models as optimizers. *arXiv preprint arXiv:2309.03409* (2023)
- S. Yao, D. Yu, J. Zhao, I. Shafran, T.L. Griffiths, Y. Cao, K. Narasimhan, Tree of thoughts: Deliberate problem solving with large language models (2023). <https://arxiv.org/pdf/2305.10601.pdf>
- D. Zhan, Y. Cheng, J. Liu, Expected improvement matrix-based infill criteria for expensive multiobjective optimization. *IEEE Trans. Evol. Comput.* **21**(6), 956–975 (2017)
- R. Zhang, D. Golovin, Random hypervolume scalarizations for provable multi-objective black box optimization, in *International Conference on Machine Learning* (PMLR, New York, 2020), pp. 11096–11105
- T. Zhang, X. Wang, D. Zhou, D. Schuurmans, J.E. Gonzalez, Tempera: test-time prompting via reinforcement learning. *arXiv preprint arXiv:2211.11890* (2022)

- G. Zhao, R. Arroyave, X. Qian, Fast exact computation of expected hypervolume improvement. *arXiv preprint arXiv:1812.07692* (2018)
- Z. Zhong, D. Friedman, D. Chen, Factual probing is [mask]: learning vs. learning to recall. *arXiv preprint arXiv:2104.05240* (2021)
- D. Zhou, L. Li, Q. Gu, Neural contextual bandits with ucb-based exploration, in *International Conference on Machine Learning* (PMLR, New York, 2020), pp. 11492–11502
- Y. Zhou, A.I. Muresanu, Z. Han, K. Paster, S. Pitis, H. Chan, J. Ba, Large language models are human-level prompt engineers. *arXiv preprint arXiv:2211.01910* (2022)
- H. Zhu, X. Wang, Y. Jin, Federated many-task bayesian optimization. *IEEE Trans. Evol. Comput.* **28**(4), 980–993 (2023)