Distributed Thompson Sampling for Large-scale Accelerated Exploration of Chemical Space

José Miguel Hernández-Lobato
University of Cambridge

Edward Pyzer-Knapp
Harvard University

Alan Aspuru-Guzik
Harvard University

Ryan P. Adams
Harvard University & Twitter

Abstract

Chemical space is so large that brute force searches for new interesting molecules are infeasible. *High-throughput virtual screening* can speed up the discovery process by collecting very large amounts of data in parallel, e.g., up to hundreds or thousands of parallel measurements. Bayesian optimization (BO) can produce additional acceleration by sequentially identifying the most useful simulations or experiments to be performed. However, current BO methods cannot scale to the large numbers of parallel measurements and the massive libraries of molecules currently used in high-throughput screening. Here, we propose a scalable solution based on a distributed implementation of the Thompson sampling heuristic. We show that this method outperforms other scalable approaches such as a greedy search strategy, ϵ-greedy approaches and a random search method.

1 Introduction

Chemical space is huge: it is estimated to contain over 10^{60} molecules. Among these, fewer than 100 million compounds can be found in public repositories or databases [20]. This discrepancy between *known* compounds and *possible* compounds implies that many molecules with useful medical or engineering functions are still to be discovered (e.g., new energy materials, pharmaceuticals, dyes, etc.). While the vast size of chemical space makes this an enormous opportunity, it also presents a significant challenge to the identification of new relevant compounds among the many unimportant ones. This challenge is so great that any discovery process relying purely on the combination of scientific intuition with trial and error experimentation is slow, tedious, and in many cases infeasible.

To accelerate the search, high-throughput approaches can be used in a combinatorial exploration of small specific areas of chemical space [18]. These have led to the development of high-throughput virtual screening [17, 6], in which large libraries of molecules are analyzed using theoretical and computational techniques before being culled to a small set of promising compounds for experimentalists to evaluate. However, even searching a tiny drop in the ocean of chemical space can result in a massive initial library of candidate compounds that hits the limits of modern computational capabilities. There is therefore an urgent need to accelerate high-throughput screening approaches.

Bayesian optimization (BO) [11] can speed up the discovery process by using machine learning models to guide the search and make improved decisions about what molecules to analyze next given the data collected so far. However, current BO methods cannot scale to the large numbers of parallel measurements and the massive libraries of candidate molecules currently used in high-throughput screening [17]. Here, we propose a scalable solution based on a distributed implementation of the Thompson sampling heuristic. We show that this method outperforms other scalable approaches such as a greedy search strategy, ϵ-greedy approaches and a random search method.

30th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain.

2 Bayesian neural networks and probabilistic back-propagation

Neural networks are well-suited for implementing BO on molecules. They produce state-of-the-art predictions of chemical properties [13, 15, 19] and can be applied to large data sets by using stochastic optimization techniques [3]. Typical applications of neural networks focus on the deterministic prediction scenario. However, when navigating chemical space, it is desirable to use a probabilistic approach that can produce accurate estimates of uncertainty for efficient exploration, and so we use *probabilistic back-propagation* (PBP), a recently-developed technique for the scalable training of Bayesian neural networks [10]. Note that other methods for approximate inference in neural networks could have been chosen as well [2, 24, 5]. We prefer PBP because it is fast and it does not require the tuning of hyper-parameters such as learning rates or regularization constants [10].

3 Distributed Thompson sampling

Thompson sampling (TS) [26] trades off exploration and exploitation automatically, without having to use additional explicit parameters, unlike ϵ-greedy approaches. TS also has the advantage that it can be easily used in the batch setting where multiple measurements are collected simultaneously. Furthermore, the process for generating each new batch of molecules can be implemented in a distributed way across different nodes in a computer cluster, unlike other methods for batched BO optimization [23, 22, 7]. This parallelization is important for speeding up computations, especially when the batch size is large and the library of candidate molecules contains millions of elements. The following lines describe the implementation of TS for collecting data in parallel using batches with N molecules:

1. A small number of molecules are selected uniformly at random from the library of candidate molecules and measurements for their ground truth values are obtained.
2. A Bayesian neural network is trained with PBP on the data collected so far.
3. The network weights are sampled N times from the posterior approximation computed by PBP. This results in N deterministic neural networks with known weight values. Each of these networks makes predictions on all the remaining candidate molecules. For each deterministic network, we select the set of N molecules with highest predicted scores. This results in N sets with N molecules each.
4. A new batch with N molecules is generated. For this, we iterate over the N sets with N molecules obtained in the previous step and select from each set the molecule with highest ranking in that set that has not been selected before.
5. Data are collected for each of the N molecules in the batch.
6. Steps 2-5 are repeated.

By sampling the network weights in step 3, TS produces N deterministic neural networks whose predictions attain a balance between exploration and exploitation. Exploitation is enforced because each deterministic neural network is an estimate of Bayes optimal predictor. Exploration is obtained as a result of the randomness of the sampling process. Thus, a balance in the search for optimal molecules is achieved automatically. Note that step 3 is a bottleneck when N and the size of the library of candidate molecules are very large. Nevertheless, this step can be easily implemented in a distributed manner by making predictions with each of the N deterministic neural networks on a different node in a computer cluster. Finally, the reason for generating N sets with N molecules in step 3 is to guarantee that we obtain a final batch with N unique molecules, even if the N deterministic networks make exactly the same predictions on the set of candidate molecules.

4 Description of the data sets

Here we describe the data sets used in our experiments. The input features for all molecules are 512-bit Morgan circular fingerprints [21], calculated with a bond radius of 2, and derived from the canonical SMILES as implemented in the RDkit package [12].

Harvard Clean Energy Project:The Clean Energy Project is the world's largest materials high-throughput virtual screening effort [9, 8], and has scanned more than 3.5 million molecules to find

those with high power conversion efficiency (PCE) using quantum-chemical techniques, taking over 30,000 years of CPU time. The target value within this data set is the power conversion efficiency (PCE), which is calculated for the 2.3 million publicly released molecules, using the Scharber model [4] and frontier orbitals calculated at the BP86 [16, 1] def2-SVP [28] level of theory.

Dose-Response Data Set: These data sets were obtained from the NCI-cancer database [14]. The dose-response target value has a potential range of -100 to 100, and reports a percentage cell growth relative to a no-drug control. Thus, a value of +40 would correspond to a 60% growth inhibition and a value of -40 would correspond to 40% lethality. Molecules with a positive value for the dose-response are known as inhibitors, molecules with a score less than 0 have a cytotoxic effect. Results against the NCI-H23 cell line were taken against a constant log-concentration of -8.00M and where multiple identical conditions were present in the data an average was used for the target variables. In this data set we are interested in finding molecules with smaller values of the target variable.

Malaria Data Set: The Malaria data set was taken from the *P. falciparum* whole cell screening derived by combining the GSK TCAMS data set, the Novatis-GNF Malaria Box data set and the St Jude's Research Hospital data set, as released through the Medicines for Malaria Venture website [25]. The target variable is the EC50 value, which is defined as the concentration of the drug which gives half maximal response. Much like the Dose response data set, the focus here is on minimization: the lower the concentration, the stronger the drug.

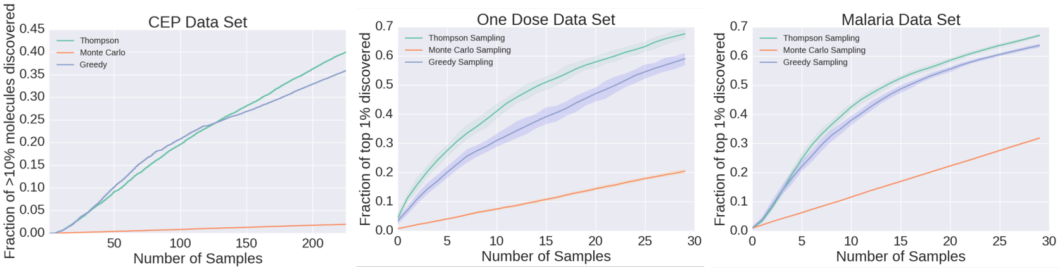

Figure 1: Recall obtained by Thompson sampling on each data set. For the CEP data, the recall for molecules with a PCE $> 10\%$ is reported, whilst for One-dose and Malaria we report the recall for the molecules in the top 1%. In addition to the Monte Carlo sampling baseline, we also include results for a greedy sampling approach, in which there is no exploration, and the molecules are chosen according to the mean of the predictive distribution given by PBP. The overall lower performance of this greedy strategy illustrates the importance of exploration when choosing molecules for high-throughput virtual screening.

5 Results

We evaluate the gains produced by Thompson sampling (TS) in experiments simulating a high throughput virtual screening setting. In these experiments, we sequentially sample molecules from libraries of candidate molecules given by the data sets from Section 4. After each sampling step, we calculate the 1% recall, that is, the fraction of the top 1% of molecules from the original library that are found among the sampled ones. For the CEP data, we compute recall by focusing on molecules with power conversion efficiency (PCE) larger than 10%. In all data sets, each sampling step involves selecting a batch of molecules among those that have not been sampled so far. In the Malaria and One-dose data sets we use batches of size 200. These data sets each contain about 20,000 molecules. By contrast, the CEP data set contains 2 million molecules. In this latter case, we use batches of size 500. We use Bayesian neural networks with one hidden layer containing 100 units.

We compare the performance of TS with two baselines. The first one, *greedy*, is a sampling strategy that only considers exploitation and does not perform any exploration. We implement this approach by selecting molecules according to the average of the probabilistic predictions generated by PBP. That is, the greedy approach ignores any variance in the predictions of the Bayesian neural network and generates batches by just ranking molecules according to the mean of the predictive distribution given by PBP. The second baseline is a Monte Carlo approach in which the batches of molecules are selected uniformly at random. These two baselines are comparable to TS in that they can be easily

Method	Rank
$\epsilon = 0.01$	3.42 ± 0.28
$\epsilon = 0.025$	3.02 ± 0.25
$\epsilon = 0.05$	2.86 ± 0.23
$\epsilon = 0.075$	3.20 ± 0.26
TS	$\mathbf{2.51 \pm 0.20}$

Table 1: Average rank and standard errors obtained by each method.

implemented in a large scale setting in which the library of candidate molecules contains millions of elements and data is sampled using large batch sizes.

In the Malaria and One-dose data sets, we average across 50 different realizations of the experiments. This is not possible in the CEP data set, which is 100 times larger than the two other data sets. In the CEP case, we report results for a single realization of the experiment (in a second realization we obtained similar results). Figure 1 shows the recall obtained by each method in the data sets from Section 4. TS significantly outperforms the Monte Carlo approach, and also offers better performance than greedy sampling. This shows the importance of building in exploration into the sampling strategy, rather than relying on purely exploitative methods. The greedy approach performs best in the CEP data set. In this case, the greedy strategy initially finds better molecules than TS, but after a while TS overtakes, probably because a promising area of chemical space initially discovered by the greedy approach starts to become exhausted.

The previous results allow us to consider the savings produced by BO. In the CEP data set, TS achieves about 20 times higher recall values than the Monte Carlo approach, which is comparable to the exhaustive enumeration that was used to collect the CEP data. We estimate that, with BO, the CEP virtual screening process would have taken 1,500 CPU years instead of the 30,000 that were actually used. Regarding the One-dose and Malaria data sets, TS can locate in both sets about 70% of the top 1% molecules by sampling approximately 6,000 molecules. By contrast, the Monte Carlo approach would require sampling 14,000 molecules. This represents a significant reduction in the discovery time for new therapeutic molecules and savings in the economic costs associated with molecule synthesis and testing.

5.1 Comparison with ϵ-greedy approaches

We can easily modify the greedy baseline from the previous section to include some amount of exploration. For example, by replacing a small fraction of the molecules in each batch with molecules chosen uniformly at random. This approach is often called ϵ-greedy [27], where the variable ϵ indicates the fraction of molecules that are sampled uniformly at random. The disadvantage of the ϵ-greedy approach is that it requires the tuning of ϵ to the problem of interest. By contrast, TS automatically adjusts the amount of exploration.

We compared TS with different versions of ϵ-greedy in the same way as above, using $\epsilon = 0.01, 0.025, 0.05$ and 0.075. The experiments with the One-dose and the Malaria data sets are similar to the ones done before. However, we now sub-sample the CEP data set to be able to average across 50 different realizations of the experiment: we choose 4,000 molecules uniformly at random and then collect data in batches of size 50 across 50 different repetitions of the screening process. We compute the average rank obtained by each method across the $3 \times 50 = 150$ simulated screening experiments. A ranking equal to 1 indicates that the method always obtains the highest recall at the end of the experiment, while a ranking equal to 5 indicates that the method always obtains the worst recall value. Table 1 shows that the lowest average rank is obtained by TS, which achieves better exploration-exploitation trade-offs than the ϵ-greedy approaches.

Acknowledgments

JMHL acknowledges support from the Rafael del Pino Foundation.

References

[1] Axel D. Becke. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics*, 98(7):5648, 1993.

[2] Charles Blundell, Julien Cornebise, Koray Kavukcuoglu, and Daan Wierstra. Weight Uncertainty in Neural Network. In *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, pages 1613–1622, 2015.

[3] Olivier Bousquet and Léon Bottou. The tradeoffs of large scale learning. In *NIPS*, pages 161–168, 2008.

[4] G. Dennler, M. C. Scharber, T. Ameri, P. Denk, K. Forberich, C. Waldauf, and C. J. Brabec. Design Rules for Donors in Bulk-Heterojunction Tandem Solar Cells?Towards 15 % Energy-Conversion Efficiency. *Adv. Mater.*, 20(3):579–583, feb 2008.

[5] Yarin Gal and Zoubin Ghahramani. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. In *Proceedings of the 33nd International Conference on Machine Learning (ICML)*, pages 1050–1059, 2016.

[6] Rafael Gómez-Bombarelli, Jorge Aguilera-Iparraguirre, Timothy D. Hirzel, David Duvenaud, Dougal Maclaurin, Martin A. Blood-Forsythe, Hyun Sik Chae, Markus Einzinger, Dong-Gwang Ha, Tony Wu, Georgios Markopoulos, Soonok Jeon, Hosuk Kang, Hiroshi Miyazaki, Masaki Numata, Sunghan Kim, Wenliang Huang, Seong Ik Hong, Marc Baldo, Ryan P. Adams, and Alán Aspuru-Guzik. Design of efficient molecular organic light-emitting diodes by a high-throughput virtual screening and experimental approach. *Nature Materials*, aug 2016.

[7] J. González, Z. Dai, P. Hennig, and N. Lawrence. Batch bayesian optimization via local penalization. In *AISTATS*, JMLR Workshop and Conference Proceedings, pages 648–657, 2016.

[8] Johannes Hachmann, Roberto Olivares-Amaya, Sule Atahan-Evrenk, Carlos Amador-Bedolla, Roel S. Sanchez-Carrera, Aryeh Gold-Parker, Leslie Vogt, Anna M. Brockway, and Alan Aspuru-Guzik. The Harvard Clean Energy Project: Large-Scale Computational Screening and Design of Organic Photovoltaics on the World Community Grid. *J. Phys. Chem. Lett.*, 2(17):2241–2251, sep 2011.

[9] Johannes Hachmann, Roberto Olivares-Amaya, Adrian Jinich, Anthony L. Appleton, Martin A. Blood-Forsythe, László R. Seress, Carolina Román-Salgado, Kai Trepte, Sule Atahan-Evrenk, Süleyman Er, Supriya Shrestha, Rajib Mondal, Anatoliy Sokolov, Zhenan Bao, and Alán Aspuru-Guzik. Lead candidates for high-performance organic photovoltaics from high-throughput quantum chemistry – the Harvard Clean Energy Project. *Energy Environ. Sci.*, 7(2):698–704, 2014.

[10] Jose Miguel Hernandez-Lobato and Ryan Adams. Probabilistic Backpropagation for Scalable Learning of Bayesian Neural Networks. In *Proceedings of The 32nd International Conference on Machine Learning*, pages 1861–1869, 2015.

[11] Donald R Jones, Matthias Schonlau, and William J Welch. Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13(4):455–492, 1998.

[12] Greg Landrum. RDKit: Open-source cheminformatics.

[13] Junshui Ma, Robert P. Sheridan, Andy Liaw, George E. Dahl, and Vladimir Svetnik. Deep Neural Nets as a Method for Quantitative Structure–Activity Relationships. *Journal of Chemical Information and Modeling*, 55(2):263–274, feb 2015.

[14] Many authors. NCI Database Download Page.

[15] Andreas Mayr, Günter Klambauer, Thomas Unterthiner, and Sepp Hochreiter. DeepTox: Toxicity Prediction using Deep Learning. *Front. Environ. Sci.*, 3, feb 2016.

[16] John P. Perdew. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B*, 33(12):8822–8824, jun 1986.

[17] Edward O. Pyzer-Knapp, Changwon Suh, Rafael Gómez-Bombarelli, Jorge Aguilera-Iparraguirre, and Alán Aspuru-Guzik. What Is High-Throughput Virtual Screening? A Perspective from Organic Materials Discovery. *Annu. Rev. Mater. Res.*, 45(1):195–216, jul 2015.

[18] Krishna Rajan. Combinatorial Materials Sciences: Experimental Strategies for Accelerated Knowledge Discovery. *Annu. Rev. Mater. Res.*, 38(1):299–322, aug 2008.

[19] Bharath Ramsundar, Steven Kearnes, Patrick Riley, Dale Webster, David Konerding, and Vijay Pande. Massively multitask networks for drug discovery. *arXiv preprint arXiv:1502.02072*, 2015.

[20] Jean-Louis Reymond, Lars Ruddigkeit, Lorenz Blum, and Ruud van Deursen. The enumeration of chemical space. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(5):717–733, apr 2012.

[21] David Rogers and Mathew Hahn. Extended-Connectivity Fingerprints. *Journal of Chemical Information and Modeling*, 50(5):742–754, may 2010.

[22] Amar Shah and Zoubin Ghahramani. Parallel predictive entropy search for batch global optimization of expensive objective functions. In *NIPS*, pages 3330–3338, 2015.

[23] Jasper Snoek, Hugo Larochelle, and Ryan P Adams. Practical bayesian optimization of machine learning algorithms. In *NIPS*, pages 2951–2959, 2012.

[24] Jasper Snoek, Oren Rippel, Kevin Swersky, Ryan Kiros, Nadathur Satish, Narayanan Sundaram, Md. Mostofa Ali Patwary, Prabhat, and Ryan P. Adams. Scalable Bayesian Optimization Using Deep Neural Networks. In *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, pages 2171–2180, 2015.

[25] Thomas Spangenberg, Jeremy N. Burrows, Paul Kowalczyk, Simon McDonald, Timothy N. C. Wells, and Paul Willis. The Open Access Malaria Box: A Drug Discovery Catalyst for Neglected Diseases. *PLoS ONE*, 8(6):e62906, jun 2013.

[26] William R. Thompson. On the Likelihood that One Unknown Probability Exceeds Another in View of the Evidence of Two Samples. *Biometrika*, 25(3/4):285, dec 1933.

[27] Christopher John Cornish Hellaby Watkins. *Learning from delayed rewards*. PhD thesis, University of Cambridge England, 1989.

[28] Florian Weigend and Reinhart Ahlrichs. Balanced basis sets of split valence triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.*, 7(18):3297, 2005.