

Scalable Optimization for Wind Farm Control using Coordination Graphs

Timothy Verstraeten
Artificial Intelligence Lab, Vrije
Universiteit Brussel
timothy.verstraeten@vub.be

Pieter-Jan Daems
Acoustics & Vibrations Research
Group, Vrije Universiteit Brussel
pieter-jan.daems@vub.be

Eugenio Bargiacchi
Artificial Intelligence Lab, Vrije
Universiteit Brussel
eugenio.bargiacchi@vub.be

Diederik M. Roijers
Vrije Universiteit Brussel
HU University of Applied Sciences
Utrecht
diederik.yamamoto-roijers@hu.nl

Pieter J.K. Libin
Data Science Institute, Hasselt
University
pieter.libin@uhasselt.be

Jan Helsen
Acoustics & Vibrations Research
Group, Vrije Universiteit Brussel
jan.helsen@vub.be

ABSTRACT

Wind farms are a crucial driver toward the generation of ecological and renewable energy. Due to their rapid increase in capacity, contemporary wind farms need to adhere to strict constraints on power output to ensure stability of the electricity grid. Specifically, a wind farm controller is required to match the farm's power production with a power demand imposed by the grid operator. This is a non-trivial optimization problem, as complex dependencies exist between the wind turbines. State-of-the-art wind farm control typically relies on physics-based heuristics that fail to capture the full load spectrum that defines a turbine's health status. When this is not taken into account, the long-term viability of the farm's turbines is put at risk. Given the complex dependencies that determine a turbine's lifetime, learning a flexible and optimal control strategy requires a data-driven approach. However, as wind farms are large-scale multi-agent systems, optimizing control strategies over the full joint action space is intractable. We propose a new learning method for wind farm control that leverages the sparse wind farm structure to factorize the optimization problem. Using a Bayesian approach, based on multi-agent Thompson sampling, we explore the factored joint action space for configurations that match the demand, while considering the lifetime of turbines. We apply our method to a grid-like wind farm layout, and evaluate configurations using a state-of-the-art wind flow simulator. Our results are competitive with a physics-based heuristic approach in terms of demand error, while, contrary to the heuristic, our method prolongs the lifetime of high-risk turbines.

KEYWORDS

Wind farm control; Multi-agent coordination; Thompson sampling

ACM Reference Format:

Timothy Verstraeten, Pieter-Jan Daems, Eugenio Bargiacchi, Diederik M. Roijers, Pieter J.K. Libin, and Jan Helsen. 2021. Scalable Optimization for Wind Farm Control using Coordination Graphs. In *Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), Online, May 3-7, 2021, IFAAMAS*, 9 pages.

Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), U. Endriss, A. Nowé, F. Dignum, A. Lomuscio (eds.), May 3-7, 2021, Online. © 2021 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

1 INTRODUCTION

As a scientific community, we acknowledge the environmental concerns associated with fossil fuel generation. Therefore, it is necessary to shift our energy generation to renewable energy sources [29]. Wind energy is expected to play a major role, experiencing a yearly growth of about 10% globally [45]. By 2030, it is anticipated that at least 30% of the European energy demand will be covered by wind energy [13].

The rapid increase in the supply of renewable energy poses challenges with respect to the stability of the electrical grid. In contrast to conventional power plants (e.g., gas, hydro and oil), the power output of wind farms ultimately depends on environmental conditions. Due to the increase in capacity, the integration of wind energy in the electricity grid needs to comply with strict grid code requirements [3, 34].

To ensure grid stability, wind farm controllers are developed to configure farm-wide power set-points, i.e., thresholds on the power production, in order to match the power demand [4]. This power demand is imposed by the transmission system operator, i.e., the entity responsible for balancing the energy supply and demand. The development of such controllers poses important challenges, as there exist complex non-linear dependencies between wind turbines. These dependencies originate from the wake effect [15] in which upstream wind turbines reduce the available wind energy for downstream wind turbines. Additionally, when wind turbines are performing torque control, which regulates the power production by adapting the rotor speed, a higher power production typically results in increased loads on the mechanical components, which leads to a higher lifetime consumption [18]. Therefore, a careful balance between power and lifetime needs to be guaranteed [8].

The design of wind farm controllers is typically grounded in domain knowledge about patterns in the turbines' behaviors when the wake effect is present [8, 32]. For example, one can recognize that upstream turbines, with respect to the dominant wind direction, typically observe higher fatigue loads (i.e., loads that induce weakening of the components) than downstream turbines [19]. Therefore, lower set-points should be chosen for upstream turbines to reduce damage accumulation through fatigue loads, in case the available power over the farm is larger than the desired

power. While such heuristics simplify the computation of the optimal set-point allocation, they fail to capture the full complexity of the dynamically-changing multi-dimensional load spectrum (e.g., loads induced during storms). In order to develop advanced control strategies, it is necessary to consider the full load spectrum to reduce failures, which increases the reliability and sustainability of wind farms [43].

In contrast to physics-based heuristics, data-driven wind farm controllers can learn control strategies without requiring in-depth knowledge about the non-linear dependencies that exist between the turbines that make up the wind farm. An example of such a data-driven structure is proposed by van Dijk et al. [39], in which reinforcement learning techniques are used to search for the optimal rotor orientation to deflect wake away from downstream turbines. However, state-of-the-art data-driven wind farm controllers scale poorly to larger wind farms, as the number of possible configurations grows exponentially with respect to the number of wind turbines.

Therefore, we argue that a hybrid approach, combining both flexible data-driven methods and physics-based domain knowledge, is key to guarantee both optimality and scalability of wind farm controllers. We propose a new method that learns farm-wide control strategies in simulation while leveraging knowledge about wake patterns, performance statistics and load profiles. Specifically, we formalize the farm-wide dependencies caused by wake as a dependency graph and cluster wind turbines with similar load profiles together to factorize the wind farm. Using this factored representation, we propose a novel sampling method, called Set-Point Thompson Sampling (SPTS). This algorithm uses multi-agent Thompson sampling to evaluate promising control strategies using the factored representation of the wind farm, with the objective to match the power demand as well as possible, while minimizing stress on wind turbines with a low remaining useful life. Additionally, we use a Bayesian formalism, which allows the inclusion of available knowledge about the data in the form of prior belief distributions. Specifically, as the expected power production of a single turbine under no-wake conditions is readily available in the turbine’s design specifications [28], we construct a prior distribution for every set-point, centered around the associated expected power production. This guides the learning process toward sensible power productions while allowing for sufficient exploration to find the optimal set-point under wake conditions.

We start by positioning our research within related work and argue that controller optimization in large-scale wind farms is necessary in Section 2. Next, we formalize the set-point configuration problem and the objective in Section 3. Then, we provide background on Thompson sampling in Section 4, after which we describe SPTS, an efficient method to explore possible set-point configurations, in Section 5. We evaluate our method on realistic wind farm settings using an extensive set of parametrizations in Section 6. Finally, we discuss the results, highlighting the benefits and limitations of the method, and conclude with future work in Section 7.

2 RELATED WORK

Wind farm control strategies have mainly focused on power-load optimization (i.e., strategies that maximize power production and minimize fatigue load) or active power control (i.e., strategies that determine power set-points to meet the demand set by the electricity grid or an operator) [8]. In both cases, the wake effect is an important factor to consider when selecting power set-points. For power-load optimization, data-driven optimization approaches typically focus on reducing the wake effect, such as wake redirection control and axial induction control. Wake redirection control is concerned with finding a joint rotor orientation of the wind turbines in order to redirect wake from downstream turbines [39, 44]. Axial induction control is an approach to reduce the wake effect, by lowering the power set-points of upstream turbines to reduce energy extraction, and thus maintain a steady wind speed behind the turbines [14, 33].

In this work, we focus on active power control to match the wind farm’s total power output to the power demand [4]. Many heuristic approaches based on physical knowledge about the turbines and environmental conditions exist [4, 19, 32, 35]. For example, one can notice that due to the wake effect, a higher power production for upstream turbines results in lower wind speeds for downstream turbines. Therefore, the power demand can be reached using a heuristic approach, where the power contributions of downstream turbines are maximized while the power contributions of upstream turbines are minimized [32].

We argue that, similar to the power-load optimization case, data-driven optimization approaches can complement existing physics-based knowledge to improve the flexibility of active power control. Such flexibility is important, as wind farm control decisions must consider the complex multi-dimensional load profiles to improve reliability [43]. Nevertheless, it remains challenging to scale data-driven optimization methods to larger wind farms, where the optimal joint configuration exists in a high-dimensional solution space.

3 PROBLEM STATEMENT

When the transmission system operator imposes a power demand, the wind farm controller needs to configure each wind turbine to a power set-point such that the total actual power production matches the demand as closely as possible. These set-points need to be chosen without assigning high-load set-points to wind turbines with a low remaining lifetime. The lifetime of a turbine is dependent on many load factors that can lead to failure. However, the link between specific loading conditions and failure is not sufficiently understood [20, 22]. Therefore, in this work, we assume that the wind farm operator constructs a cost-function that heavily penalizes high loads on high-risk turbines based on expert knowledge (see Section 8). We formalize the setting as a tuple $\langle \mathcal{W}, G, \mathcal{R}, \mathcal{S}, \langle P, L \rangle, P_{\text{dem}} \rangle$, which can be regarded as an extension of a multi-agent multi-armed bandit [7, 36], where

- \mathcal{W} is a set of wind turbines.
- G is a directed dependency graph that describes which subset of turbines influences a particular reference turbine. We refer to the dependencies of a wind turbine $w \in \mathcal{W}$ as its *parents*,

and denote the set that comprises turbine w and its parents as $G(w)$.

- \mathcal{R} is a set of operational zones, or *regimes*, within the wind farm, in which turbines observe similar loads under normal operating conditions. A fraction of the power demand will be allocated to each of these regimes.
- $\mathcal{S} = \mathcal{S}^1 \times \dots \times \mathcal{S}^{|\mathcal{W}|}$ is the set of joint set-point configurations, which is the Cartesian product of the turbine-specific set-points $s^w \in \mathcal{S}^w$ for each turbine $w \in \mathcal{W}$. We denote $\mathcal{S}^{G(w)}$ as the set of local joint set-points for the set $G(w)$.
- $P(s)$ is a stochastic function providing the farm-wide power production when a joint set-point configuration, $s \in \mathcal{S}$, is evaluated. The global power production can be decomposed into $|\mathcal{W}|$ observable and independent local functions, i.e., $P(s) = \sum_{w \in \mathcal{W}} P^w(s^{G(w)})$. The local function $P^w(s^{G(w)})$ represents the power production achieved by wind turbine w and only depends on the local joint set-point $s^{G(w)}$ of the subset of wind turbines in $G(w)$.
- $L^w(s^{G(w)})$ assigns a penalty for performing high-load actions by wind turbine w . We assume that the wind farm operator heavily penalizes high-risk turbines (e.g., machines that have a low remaining life) based on available domain knowledge.
- P_{dem} is the power demand imposed by the transmission system operator.

We aim to find the joint set-point configuration that matches the power demand as well as possible, while penalizing high-load actions on wind turbines with low remaining useful life:

$$\min_s \sum_{r \in \mathcal{R}} \left| f_r P_{\text{dem}} - \sum_{w \in \mathcal{R}} P^w(s^{G(w)}) \right| + \sum_{w \in \mathcal{W}} L^w(s^{G(w)}), \quad (1)$$

where f_r is a parameter that assigns a fraction of the demand to regime r .

4 THOMPSON SAMPLING

Thompson sampling (TS) is a decision making strategy, in which the exploration process for the optimal action is guided by the user's prior beliefs about the unknown rewards of the actions. Specifically, using a Bayesian formalism, it models the unknown parameters of the reward distribution for various pre-defined actions as prior belief distributions. The user should provide these distributions and quantify the existing uncertainty and information about these parameters, prior to learning. Once a particular action has been evaluated and its reward has been measured, TS uses Bayes' rule to update the belief distributions, and thus maintains posteriors over the unknown parameters. Given a prior Q_a for action a , and a history of observations \mathcal{H}_{t-1} , the user's beliefs about the expected reward $\mu(a)$ at time t is modeled as:

$$\mu(a) \sim Q_a(\cdot | \mathcal{H}_{t-1}). \quad (2)$$

TS is a probability matching mechanism [24], as it directly samples an action according to the probability that it is optimal at time step t . Specifically, for each action, it draws a sample from the associated posterior distribution, representing the expected reward

of the action. Afterwards, it computes the action with the highest expected reward over all samples. Formally,

$$\begin{aligned} \mu_t(a) &\sim Q_a(\cdot | \mathcal{H}_{t-1}), \forall a \\ a_t &= \arg \max_a \mu_t(a). \end{aligned} \quad (3)$$

Finally, the chosen action is evaluated, and its measured reward (which may be stochastic) is added to the history of observations.

Note that the chosen action is independent and identically distributed with respect to the optimal action according to the user's beliefs [24], i.e.,

$$p(a_t | \mathcal{H}_{t-1}) = p(a_* | \mathcal{H}_{t-1}), \quad (4)$$

where a_* is the optimal action. Thus, TS balances exploration and exploitation by focusing on promising actions, while considering the user's uncertainty about the problem. By allowing the user to configure prior belief about the problem, TS typically performs well in practical applications, where such information is often readily available [10, 25, 26].

In multi-agent settings, the joint action space scales exponentially with the number of agents [11]. As TS does not leverage a factored representation of the joint action space, it becomes intractable to use for large multi-agent systems. Therefore, it is important to exploit the structure of the multi-agent setting whenever available, in order to decrease the complexity of the optimization problem [7, 23].

5 SET-POINT THOMPSON SAMPLING

To ensure scalability of the control optimization process toward large wind farms, we need a learning algorithm that can decompose and exploit the farm's topology. To this end, we propose a new wind farm control algorithm, Set-Point Thompson Sampling (SPTS), which constructs a dependency graph and regimes from wake fields and load profiles, and relies on multi-agent Thompson sampling (MATS) [40, 41] to efficiently explore the joint set-point configuration space using the factored representation.

5.1 Factorization

To accurately decompose the problem, we rely on two aspects. Firstly, wind turbines depend on each other due to the wake effect. The decisions made by upstream turbines affect downstream turbines. Therefore, local coordination between a reference turbine and its affected neighbors is necessary to guarantee optimality of the solution. Secondly, due to the wake effect, upstream wind turbines produce more power than downstream turbines. As power and torque loading are highly correlated, wind turbines with similar power productions overall have similar fatigue load profiles [5, 43]. Since turbines with a lower observed fatigue load should be responsible for the majority of the demand, a group of similar turbines can be assigned a fraction of the demand inversely proportional to their observed loads.

To construct the dependency graph G , we analyze the wake field generated at a given wind direction. Using a geometric approach, we derive, for a particular reference turbine, which downstream turbines are in its wake. Specifically, we consider a downstream turbine w to be dependent on an upstream reference turbine w_{ref} if and only if w is geographically positioned within a specified

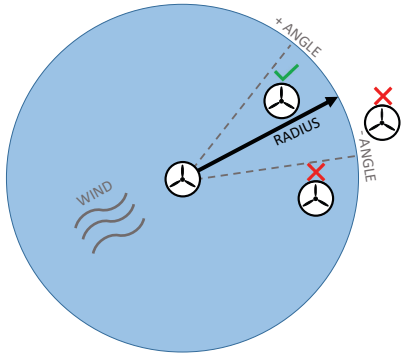


Figure 1: An example with one upstream reference turbine and three downstream turbines. Only one downstream turbine is considered to be dependent on the reference turbine (green check mark), as it is both within the specified radius and within the specified angle with respect to the incoming wind vector from the reference turbine.

radius from w_{ref} and is within a specified angle from the incoming wind vector. Naturally, a larger angle and radius would lead to a more dense dependency graph. For our purpose, we found that downstream turbines located at an angle of 22.5° and a radius of 1 km sufficiently lie outside of the wake effect induced by the reference turbine, according to the simple Jensen wake model [21]. A visualisation of the geometric approach is shown in Figure 1. Examples of a wake field and corresponding dependency graph are shown in Figure 2.

To construct the operational regimes \mathcal{R} , we group wind turbines based on their observed fatigue loads. As mentioned before, turbines with similar overall power productions have similar load profiles. Since the majority of turbine operations are performed in steady states under the dominant wind direction, we cluster the turbines based on their power productions under active wake conditions. Specifically, we cluster the turbines using the state-of-the-art approach described in [43], which uses a Bayesian Gaussian mixture model. These models describe the data using a finite set of multivariate Gaussian distributions. Afterwards, each turbine is assigned to the regime (i.e., Gaussian) it most likely belongs to.

Once the regimes have been created, we assign a fraction f_r of the demand to each regime r . This fraction should be proportional to the remaining useful life of the regime. We construct the load revolution distribution [18], i.e., the number of rotations performed by the turbines within a regime operating at a particular torque level, for each operational regime, based on real wind farm data of 24 turbines. We define the remaining lifetime in terms of the accumulated damage, which is the number of rotations operated under high-load conditions. We consider an operation to be high-load when the expected (main shaft) torque exceeds a particular threshold. Formally, the torque of turbine w at time t is defined as

$$\tau_t^w = \frac{P_t^w}{\omega_t^w}, \quad (5)$$

where P_t^w and ω_t^w are, respectively, the produced power (in W) and angular rotor speed (in rad/s) of turbine w at time t . We assume

that damage occurs when the torque exceeds the one achieved at 65% of the turbine’s maximum power level [5]. Thus, we can derive the number of rotations performed at a torque higher than this threshold from the load revolution distribution of each turbine and aggregate the results per regime. After normalization of the remaining rotations over the entire wind farm, the fraction f_r is set to the inverse of the sum over all turbines within regime r . An example of regimes, associated with the normalized number of life-consuming rotations per turbine, is shown in Figure 3.

5.2 Exploration

Consider the problem described in Section 3. The expected power productions $P^w(s^{G(w)})$ for each possible local set-point configuration $s^{G(w)}$ are unknown. Similar to MATS [40], SPTS uses a Bayesian formalism, which means users can exert their beliefs over $P^w(s^{G(w)})$ in the form of a prior. If wind turbines were not affected by wake, the incoming wind speed can be used to predict the expected power production of the wind turbine, which is provided with the turbine design specifications [28]. Therefore, for a given set-point, wind speed and turbine, we model the achieved power using a Gaussian prior, where the mean is the expected power production given the set-point under no wake conditions, and the standard deviation σ represents prior uncertainty about the achieved power.

$$P^w(s) \sim \mathcal{N}(\cdot \mid \mu_{s^w}^w, \sigma), \quad (6)$$

$$\mu_{s^w}^w = \min(s^w, P_{\text{av}}^w),$$

where the mean is the minimum between the set-point s^w and the available power P_{av}^w . The standard deviation σ balances exploration of alternative power production outcomes (high σ), and exploitation of the provided domain knowledge (low σ).

At each time step t , SPTS draws a sample $P_t^w(s^{G(w)})$ from the posterior for each wind turbine and possible local set-point configuration, given the history, \mathcal{H}_{t-1} , consisting of previously evaluated set-points and associated power productions:

$$P_t^w(s^{G(w)}) \sim \mathcal{N}(\cdot \mid \mu_{s^w}^w, \sigma, s^{G(w)}, \mathcal{H}_{t-1}), \text{ with} \quad (7)$$

$$\mathcal{H}_{t-1} = \bigcup_{i=1}^{t-1} \bigcup_{w \in \mathcal{W}} \left\{ \left\langle s_i^{G(w)}, P^w(s_i^{G(w)}) \right\rangle \right\}.$$

Note that during this step, SPTS samples directly the posterior over the unknown local means, which implies that the sample $P_t^w(s^{G(w)})$ and the unknown mean $P^w(s^{G(w)})$ are independent and identically distributed at time step t , given history \mathcal{H}_{t-1} .

SPTS takes the set-point that minimizes the objective function (see Equation 1). In traditional Thompson sampling (TS) [38], the optimal solution is found by maximizing over the full joint action space. However, this is intractable for larger multi-agent settings, as the joint action space scales exponentially with the number of agents. For example, a moderately-sized wind farm comprised of 20 wind turbines, where each turbine can choose from 3 possible set-points, would have 3^{20} (approximately 3.5 billion) possible configurations. Fortunately, due to the structure of wind farms, the optimal set-point configuration exists in the sparse factored

representation of the joint action space. Therefore, the minimization problem defined in Equation 1 can be solved exactly and effectively using variable elimination [17] or linear programming [27].

Finally, the joint set-point configuration that minimizes the objective function, s_t , is executed in simulation and the associated power productions $P_t^w(s_t^{G(w)})$ will be recorded for each wind turbine w . SPTS is formally described in Algorithm 1.

6 EXPERIMENTS

As the majority of offshore wind farms have a symmetric grid-like topology [37], we conduct our experiments in a wind farm that has the shape of a parallelogram. Grid-like layouts are often beneficial toward the planning and construction of the farms. However, such layouts cause wake due to the proximity of the turbines, reducing the overall power production of the wind farm [37]. We place 24 turbines in a 4-by-6 grid, 500 m apart along the x-axis and 400 m apart along the y-axis, as shown in Figure 2.

We investigate our method under the same global wind conditions used by [32], in which we assume that 0° is the dominant wind direction and orthogonal to the wind farm grid. Specifically, we investigate the incoming wind vectors at 0° and 30° , with a speed of 11 m/s. Under these conditions, the wake effect is strong. The wake fields and dependency graphs for both wind vectors are shown in Figure 2.

We use FLORIS [30], a state-of-the-art simulator (constructed by the National Renewable Energy Lab in the USA) to simulate farm and wake conditions, and the LW-8MW reference turbine [12] to model the turbines' dynamics. These models accurately represent the scale of, and conditions at, contemporary offshore wind farms. From real wind farm data, we compute damage-inducing fatigue

loads for every turbine, according to [5], and derive the fractions f_r as discussed in Section 5. The normalized damage-inducing fatigue loads are reported in Figure 3. We provide 3 possible set-points to a wind turbine w , i.e., $s_w \in \{1490 \text{ kW}, 6420 \text{ kW}, 8000 \text{ kW}\}$ (equivalent to measured wind speeds of 6.5, 10.0 or 13.5 m/s at the turbine's location), which translates into a low, medium and high power production.

We perform experiments for all combinations of the following sets of parameters:

- Wind direction: $d \in \{0^\circ, 30^\circ\}$
- Demand: $P_{\text{dem}} \in \{60 \text{ MW}, 70 \text{ MW}, 80 \text{ MW}, 90 \text{ MW}, 100 \text{ MW}\}$
- Number of high-risk turbines: $n \in \{1, 2, 3, 4\}$
- Penalty: $L^w(s^{G(w)}) = \begin{cases} +\infty & \text{if } w \text{ is high-risk and} \\ & P_t^w(s^{G(w)}) \geq 5.2 \text{ MW} \\ 0 & \text{otherwise} \end{cases}$

An infinite penalty is provided to a high-risk turbine when the used set-point leads to damage accumulation, which occurs at 65% of the maximum power level [5]. Although any non-linear penalty function can be used, this infinite penalty allows the wind farm operator to identify high-risk turbines and ensure that no high-load set-points are assigned to them (see Section 8). The n high-risk turbines are randomly chosen. We set the standard deviation σ in the prior distribution (Equation 6) to 1 MW, which allows for a sufficient amount of exploration over the complete power range of [0 MW, 8 MW]. Each experiment is repeated 100 times.¹

We compare SPTS with a set-point allocation strategy, based on the heuristic proposed by [32]. Specifically, higher set-points are first assigned to turbines which are further back in the farm, with respect to the incoming wind vector. This process is repeated toward the front of the farm until the required demand is reached. The solution with the power production that is closest to the demand is recorded. For both the heuristic and SPTS, the best performing control strategies of each run are compared. As the main focus is to prevent high-load actions on high-risk turbines, we define the performance of a set-point configuration in terms of the total penalty first, and in case of draws, the configuration that matches the demand the closest is chosen.

Figure 4 shows the learning curve of SPTS for the setting with a wind direction of 0° , a demand of 80 MW and 3 high-risk turbines. The trend indicates that 200 iterations are sufficient to ensure convergence. The learning curves for all settings are reported in the Supplementary Material.² Note that the heuristic is a deterministic approach, and thus the variance on the outcomes, over multiple repetitions of the experiment, is zero.

Figure 5 shows the average absolute difference between the best performances of the heuristic and of SPTS for all parameter combinations, both in terms of penalty and demand error. On the one hand, SPTS reaches comparable or better results than the heuristic in terms of demand error. On the other hand, the heuristic approach receives more penalties when the number of high-risk turbines is increased, or when the required demand is increased.

Algorithm 1: SPTS

```

1  $G, \mathcal{R} \leftarrow$  Construct dependency graph and regimes
2  $\mathcal{H}_0 \leftarrow \{\}$ 
3 for  $t \in [1..T]$  do
4   Sample expected performance for every possible local
   set-point configuration.
5   for  $w \in \mathcal{W}, s \in \mathcal{S}^{G(w)}$  do
6      $P^w(s) \sim \mathcal{N}(\cdot | \mu_{s^w}^w, \sigma, \mathcal{H}_{t-1})$ 
7   end
8   Select best joint configuration.
9    $s_t \leftarrow \arg \min_s \sum_{r \in \mathcal{R}} |f_r P_{\text{dem}} - \sum_{w \in r} P^w(s^{G(w)})|$ 
10     $+ \sum_{w \in \mathcal{W}} L^w(s^{G(w)})$ 
11   Simulate chosen set-point configuration.
12    $\langle P_t^w(s_t^{G(w)}) \rangle_{w \in \mathcal{W}} \leftarrow$  Simulate configuration  $s_t$ 
13   Update belief distributions using observed performance.
14    $\mathcal{H}_t \leftarrow \mathcal{H}_{t-1} \cup \left\{ \left\langle s_t^{G(w)}, P_t^w(s_t^{G(w)}) \right\rangle_{w \in \mathcal{W}} \right\}$ 
15 end

```

¹The source code of all experiments is publicly available at <https://github.com/timoverstraeten/spts-experiments>.

²Supplementary Material: <https://zenodo.org/record/4446701>

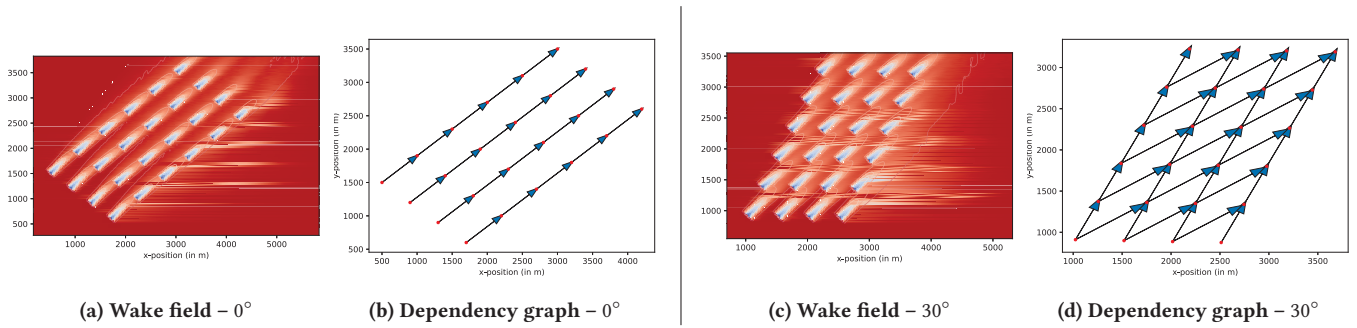


Figure 2: Wind farm – Based on the wake field generated for a particular wind direction, wind speed of 11 m/s, and maximal power set-points, we create a dependency graph. We show the results for wind directions of 0° (a+b) and 30° (c+d). For a wind direction of 30° , the wind farm is rotated by -30° , such that the global wind vector always starts at $(0, 0)$.

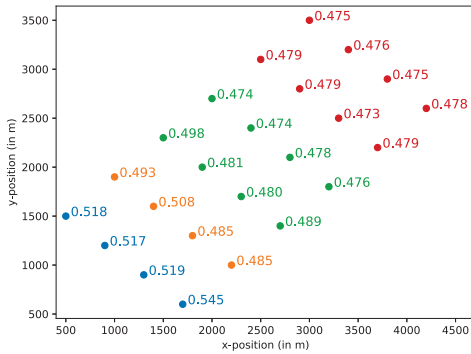


Figure 3: Operational regimes in the wind farm, obtained under the dominant wind vector coming from the origin $(0, 0)$. The turbines are annotated with their observed normalized damage-inducing load.

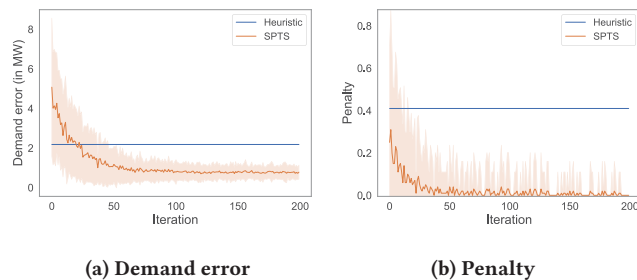


Figure 4: Learning curves of the demand error (a) and total penalty (b) obtained at each iteration, for a wind direction of 0° , a demand of 80 MW and 3 high-risk turbines. The average trend (line) and standard deviation (shaded area) are shown. The experiment is repeated 100 times.

This is expected, as the heuristic approach will allocate higher set-points to high-risk machines to reach a higher demand. Moreover, SPTS significantly outperforms the heuristic in terms of demand error when the demand is 80 MW. This is due to the fact that many

possible set-point configurations exist to meet this demand, and can thus easily be found by SPTS. In contrast, when the demand is 60 MW or 100 MW, only a few configurations are viable, in which most of the set-points are low or high, respectively. It is important to note that, over all runs, the best configurations achieved by SPTS *never* contained a damage-inducing set-point assigned to a high-risk turbine (i.e., the total penalty of the solution is always 0 for all settings).

Box plots of the best performing control strategies obtained by SPTS and the heuristic for all parameter combinations are reported in the Supplementary Material.

Figure 6 provides a farm-wide view of the power productions obtained, given the best set-point configuration found by the heuristic approach and SPTS. This is shown for the setting with a wind direction of 0° , a demand of 80 MW and 3 high-risk turbines. The farm-wide power productions for all settings are reported in the Supplementary Material. By the heuristic's definition, set-points are first maximized for downstream turbines. This process is repeated toward the front of the farm. Therefore, it is expected to see a pattern in which the downstream turbines exhibit a higher actual power production, compared to the upstream turbines. While the heuristic generates an intuitive pattern, the solution of SPTS indicates that this is not necessarily an optimal solution. Therefore, the results shown in Figure 5a and 6 demonstrate that a data-driven approach is essential in order to come up with context-aware solutions. Nevertheless, due to the allocation of demand over the different regimes, SPTS also favors higher set-points for downstream turbines. The reduced power generation of turbines with higher observed loads, combined with the ability to include arbitrary penalty functions (see Figure 5b), demonstrates that SPTS is suitable to match the power demand in a manner that reduces overall lifetime consumption.

7 DISCUSSION

We propose a new wind farm control algorithm that allocates set-points taking into account load information. The results demonstrate that SPTS can successfully match the power demand without assigning high-load set-points to turbines.

We focus on a wind farm with a symmetric grid-like structure, which comprises the majority of contemporary wind farm designs.

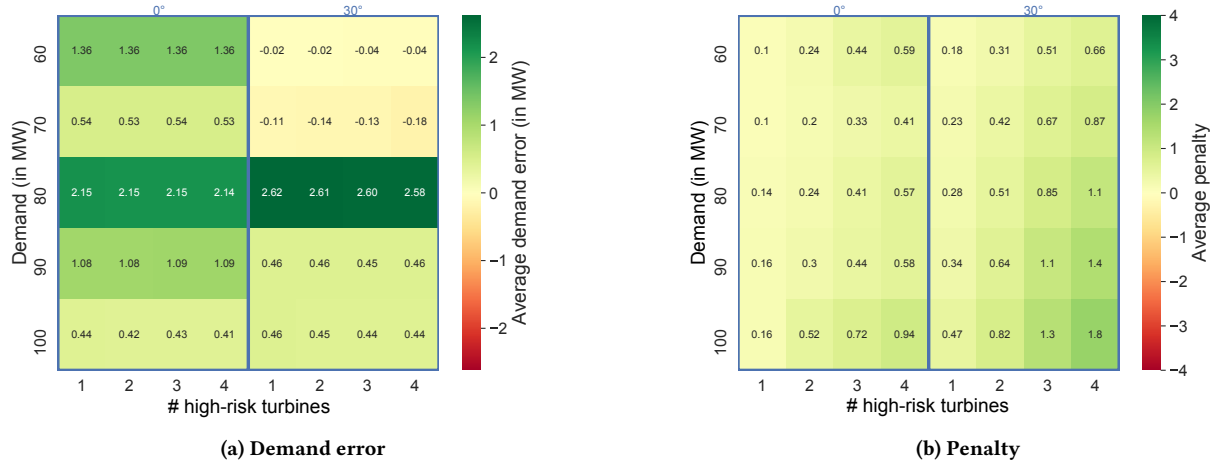


Figure 5: Heatmaps of the average absolute difference between the performance of the heuristic and of SPTS. The performance in terms of demand error (left) and total penalty (right) are plotted. A positive value (green) indicates a better average performance obtained by SPTS, while a negative value (red) indicates a worse performance compared to the heuristic. The performances are averaged over 100 samples.

However, our method is not biased toward symmetric topologies. Instead, SPTS guides the learning process by using information about the environmental and operational conditions at the farm site, acquired from wake analyses and knowledge regarding the turbines’ health. Therefore, SPTS is applicable to irregularly shaped wind farms, which can be constructed due to landscape constraints [16].

To model the turbine’s dynamics and wind flow, we use the FLORIS wake simulator, in which set-point configurations can be evaluated quickly. SPTS has the ability to include expert knowledge without being dependent on the simulator used. Therefore, it is easy to switch between different types of simulators, such as computational fluid dynamics simulators [9].

In our experiments, we focused on optimizing control strategies under steady wind conditions. Therefore, it is sufficient to analyze the expected performance of control strategies in a noise-free setting. Nevertheless, in the case of transient wind conditions (e.g., wind gusts or storms), it may be important to investigate the variance on the performance measure as well. Since SPTS uses a Bayesian sampling approach, it is straightforward to introduce a likelihood distribution with a (possibly unknown) noise parameter and update the posterior distributions according to Bayes’ rule [31].

Our approach uses multi-agent Thompson sampling (MATS) for sampling the set-point configuration space. For MATS, an asymptotic upper bound was established on the cumulative regret, i.e., the total performance loss obtained by executing sub-optimal actions during the learning phase. As we investigated a noise-free setting, every set-point configuration will likely be executed once eventually, and thus the optimal set-point configuration will be found. Therefore, the asymptotic upper bound has no practical use here. Nevertheless, the composition of the bound suggests that the performance loss of SPTS is in terms of the number of local joint set-points, rather than the number of global joint set-points, as is

the case in traditional TS. Therefore, it is expected that SPTS can learn efficiently using noisy set-point evaluations as well.

Finally, SPTS is a sampling technique inspired by TS to select promising joint actions. Recently, theoretical guarantees have been established for TS in terms of cumulative regret, i.e., the total difference between the expected reward of the optimal (unknown) action and chosen actions obtained *during* the learning process [1]. However, as the set-point optimization is performed *in silico*, the performance of the evaluated set-points during the learning phase is only used to guide the learning process. In our setting, we focus on the performance of the end result, i.e., the best set-point configuration obtained *after* learning. This highlights the distinction between exploration-exploitation and best arm identification [6]. Note that this is a difference in terms of convergence speed, rather than optimality, as the action chosen by TS still converges to the optimal one (under certain conditions) [24]. To our knowledge, there are no Bayesian best arm identification algorithms available for dealing with factored multi-agent systems. Although the learning curves converge quickly (see Supplementary Material), further research into best arm identification algorithms for loosely-coupled multi-agents is warranted to improve sample efficiency.

8 FUTURE CHALLENGES

In our experiments, we use objective functions that heavily penalize a set of turbines, which allows wind farm operators to mark high-risk turbines and reduce their loading conditions. Still, SPTS finds the optimal joint set-point configuration under arbitrary penalties (see Section 3). Such flexibility is required for capturing the multi-dimensional load spectrum that is present in wind turbine technology. However, further research needs to be conducted to establish the links between turbine responses during dynamic events and potential failure modes [20, 22, 43].

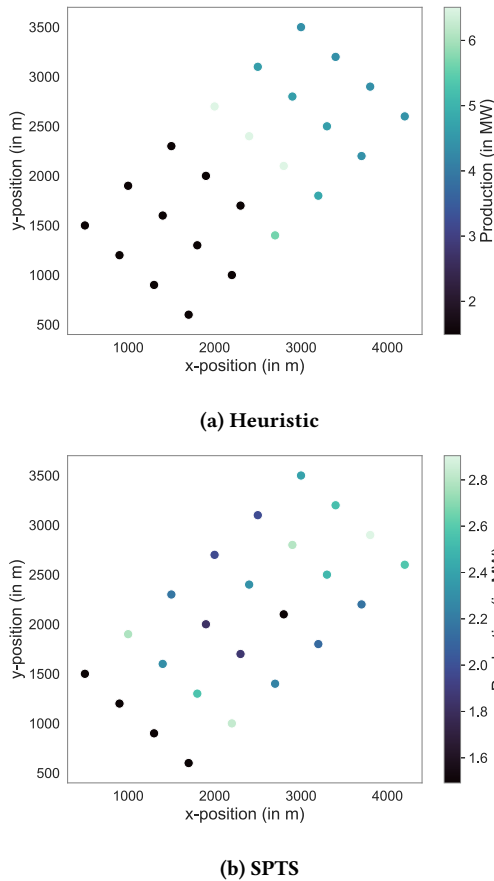


Figure 6: Farm-wide view of the power productions for the wind direction of 0° , a demand of 80 MW and 3 high-risk turbines. The power productions for the best set-point configurations found by the heuristic approach (a) and SPTS (b) are shown. The power productions are averaged over 100 runs.

Through fundamental research and condition monitoring analyses, maintenance costs can be formalized as a penalty function within SPTS, which is important toward the further development of advanced wind farm controllers.

SPTS finds the optimal combination of set-points taken from a discrete space. To further reduce the demand error, continuous set-points should be considered. Optimizing over a factored representation in continuous space is challenging, as the choice of a turbine is possibly dependent on an infinite amount of configurations chosen by its parents. To our knowledge, no optimization algorithms currently exist that fully operate within a factored continuous action space. Therefore, research should focus on continuous optimization techniques for loosely-coupled multi-agent systems, such that accurate solutions can be provided in a feasible manner. Nevertheless, one can use the optimal discrete set-point configuration provided by SPTS as a starting point and further optimize over the continuous set-point space using an iterative

approach [32]. Naturally, such a solution may not be provably optimal and convergence may be challenging to guarantee.

Our approach considers each scenario independently and does not generalize over environmental conditions. In data-driven wind farm control research, control strategies are often learned without considering the dependencies between environmental parameters (e.g., wind speed and wind direction) [39, 43]. However, generalizing over environmental conditions, for instance, through the use of contextual bandits [2], would improve the sample-efficiency of SPTS over multiple settings.

Finally, to accurately represent the current environmental conditions at the real wind farm, a mechanism is necessary to reduce measurement noise, while retaining the fine granularity of the data sources. To this end, a similarity-based data exchange between the turbines can be used to process real wind farm data [42] and provide a reliable parametrization to the wake simulator. Combined with the SPTS control algorithm, scenario optimization can be performed to provide set-point configurations in real time, allowing SPTS to be used in the real world.

9 CONCLUSION

We proposed Set-Point Thompson sampling (SPTS), a novel Bayesian algorithm for wind farm control. The method exploits the sparse topological structure of wind farms to efficiently search for the optimal set-point configuration. We showed that SPTS achieves a comparable or better performance than the heuristic approach in terms of demand error. Yet, SPTS manages to consistently assign lower set-points to high-risk turbines, which effectively reduces their lifetime consumption.

The results demonstrate that a data-driven approach is necessary when dealing with complex penalty functions. Complex penalty functions are inevitable when considering the multi-dimensional load spectrum of wind turbine technology. Introducing knowledge about the load spectrum is key to advance the current state-of-the-art in wind farm control, as loads have a large impact on the remaining useful life of the turbines, and ultimately on maintenance costs.

Still, the ability to introduce wind farm properties and mechanical knowledge into the data-driven wind farm controller is necessary to render the learning process tractable. Therefore, closing the gap between physics-based heuristics and data-driven learning will ensure the development of advanced wind farm controllers that are well-balanced in terms of learning tractability, flexibility and optimality.

ACKNOWLEDGMENTS

The authors would like to acknowledge FWO (Fonds Wetenschappelijk Onderzoek) for their support through the SB grants of Timothy Verstraeten (#1S47617N) and Eugenio Bargiacchi (#1SA2820N), and the post-doctoral grant of Pieter J.K. Libin (#1242021N). This research was supported by funding from the Flemish Government under the ‘‘Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen’’ programme and under the VLAIO Supersized 4.0 ICON project.

REFERENCES

- [1] Shipra Agrawal and Navin Goyal. 2012. Analysis of Thompson sampling for the multi-armed bandit problem. In *Proceedings of the 25th Annual Conference on Learning Theory (COLT)*, Vol. 23. 39.1–39.26.
- [2] Shipra Agrawal and Navin Goyal. 2013. Thompson sampling for contextual bandits with linear payoffs. In *International Conference on Machine Learning (ICML)*. 127–135.
- [3] Shakir D Ahmed, Fahad SM Al-Ismael, Md Shafiqullah, Fahad A Al-Sulaiman, and Ibrahim M El-Amin. 2020. Grid integration challenges of wind energy: A review. *IEEE Access* 8 (2020), 10857–10878.
- [4] J. Aho, A. Buckspan, J. Laks, P. Fleming, Y. Jeong, F. Dunne, M. Churchfield, L. Pao, and K. Johnson. 2012. A tutorial of wind turbine control for supporting grid frequency through active power control. In *American Control Conference (ACC)*. 3120–3131.
- [5] Eduardo J Alvarez and Adrijan P Ribaric. 2018. An improved-accuracy method for fatigue load analysis of wind turbine gearbox based on SCADA. *Renewable Energy* 115 (2018), 391–399.
- [6] Jean-Yves Audibert and Sébastien Bubeck. 2010. Best arm identification in multi-armed bandits. In *Proc. of the 23rd Annual Conference on Learning Theory (COLT)*.
- [7] Eugenio Bargiacchi, Timothy Verstraeten, Diederik Roijers, Ann Nowé, and Hado van Hasselt. 2018. Learning to Coordinate with Coordination Graphs in Repeated Single-Stage Multi-Agent Decision Problems. In *International Conference on Machine Learning (ICML)*. 491–499.
- [8] Sjoerd Boersma, BM Doekemeijer, Pieter MO Gebraad, Paul A Fleming, Jennifer Annoni, Andrew K Scholbrock, JA Frederik, and Jan-Willem van Wingerden. 2017. A tutorial on control-oriented modeling and control of wind farms. In *2017 American Control Conference (ACC)*. IEEE, 1–18.
- [9] Francesco Castellani, Arne Gravdahl, Giorgio Crasto, Emanuele Piccioni, and Andrea Vignaroli. 2013. A practical approach in the CFD simulation of off-shore wind farms through the actuator disc technique. *Energy Procedia* 35 (2013), 274–284.
- [10] Olivier Chapelle and Lihong Li. 2011. An empirical evaluation of Thompson sampling. In *Advances in Neural Information Processing Systems (NIPS)*, Vol. 24. 2249–2257.
- [11] Caroline Claus and Craig Boutilier. 1998. The dynamics of reinforcement learning in cooperative multiagent systems. *Proc. of the 15th National Conference on Artificial Intelligence* (1998), 746–752.
- [12] Cian Desmond, Jimmy Murphy, Lindert Blonk, and Wouter Haans. 2016. Description of an 8 MW reference wind turbine. In *Journal of Physics: Conference Series*, Vol. 753. IOP Publishing.
- [13] Wind Europe. 2017. *Wind energy in Europe: Scenarios for 2030*. Technical Report. <https://windeurope.org/about-wind/reports/wind-energy-in-europe-scenarios-for-2030/>.
- [14] Pieter MO Gebraad and Jan-Willem van Wingerden. 2015. Maximum power-point tracking control for wind farms. *Wind Energy* 18, 3 (2015), 429–447.
- [15] F González-Longatt, P Wall, and V Terzija. 2012. Wake effect in wind farm performance: Steady-state and dynamic behavior. *Renewable Energy* 39, 1 (2012), 329–338.
- [16] Huajie Gu and Jun Wang. 2013. Irregular-shape wind farm micro-siting optimization. *Energy* 57 (2013), 535–544.
- [17] Carlos Guestrin, Daphne Koller, and Ronald Parr. 2001. Multiagent Planning with Factored MDPs. In *Advances in Neural Information Processing Systems (NIPS)*, Vol. 14. 1523–1530.
- [18] International Electrotechnical Commission. 2012. Wind Turbines – Part 4: Design requirements for wind turbine gearboxes (No. IEC 61400-4).
- [19] T. Jensen, Torben Knudsen, and Thomas Bak. 2016. Fatigue minimising power reference control of a de-rated wind farm. *Journal of Physics: Conference Series* 753 (2016).
- [20] V. Jantara Junior, J. Zhou, S. Roshanmanesh, F. Hayati, S. Hajiabady, X. Y. Li, H. Dong, and M. Papaalias. 2017. Evaluation of damage mechanics of industrial wind turbine gearboxes. *Insight – Non-Destructive Testing and Condition Monitoring* 59, 8 (2017), 410–414.
- [21] I Katic, Jørgen Højstrup, and Niels Otto Jensen. 1986. A simple model for cluster efficiency. In *European wind energy association conference and exhibition*, Vol. 1. 407–410.
- [22] Jonathan Keller, S. Sheng, J. Cotrell, and A. Greco. 2016. *Wind Turbine Drivetrain Reliability Collaborative Workshop: A recap*. Technical Report. National Renewable Energy Laboratory (NREL), Golden, CO, United States.
- [23] Jelle R. Kok and Nikos Vlassis. 2004. Sparse Cooperative Q-learning. In *Proceedings of the 21st International Conference on Machine Learning (ICML)* (Banff, Alberta, Canada). New York, NY, USA.
- [24] Tor Lattimore and Csaba Szepesvári. 2020. *Bandit Algorithms*. Cambridge University Press. <https://doi.org/10.1017/9781108571401>
- [25] Pieter J. K. Libin, Timothy Verstraeten, Diederik M. Roijers, Jelena Grujic, Kristof Theys, Philippe Lemey, and Ann Nowé. 2018. Bayesian Best-Arm Identification for Selecting Influenza Mitigation Strategies. In *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD)*. Springer, 456–471.
- [26] Pieter J. K. Libin, Timothy Verstraeten, Diederik M. Roijers, Wenjia Wang, Kristof Theys, and Ann Nowé. 2019. Thompson Sampling for m-top Exploration. In *Proceedings of the IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*. 1414–1420.
- [27] Robin Lougee-Heimer. 2003. The Common Optimization Interface for Operations Research. *IBM Journal of Research and Development* 47, 1 (2003), 57–66.
- [28] M Lydia, S Suresh Kumar, A Immanuel Selvakumar, and G Edwin Prem Kumar. 2014. A comprehensive review on wind turbine power curve modeling techniques. *Renewable and Sustainable Energy Reviews* 30 (2014), 452–460.
- [29] Sourav Mukherjee, Ashok Mishra, and Kevin E. Trenberth. 2018. Climate Change and Drought: a Perspective on Drought Indices. *Current Climate Change Reports* 4, 2 (2018), 145–163.
- [30] NREL. 2019. FLORIS. Version 2.0.1. <https://github.com/NREL/floris>.
- [31] Daniel Russo, Benjamin Van Roy, Abbas Kazerouni, Ian Osband, and Zheng Wen. 2017. A tutorial on thompson sampling. *arXiv preprint arXiv:1707.02038* (2017).
- [32] Sara Siniscalchi-Minna, Fernando D Bianchi, Mikel De-Prada-Gil, and Carlos Ocampo-Martinez. 2019. A wind farm control strategy for power reserve maximization. *Renewable energy* 131 (2019), 37–44.
- [33] Maryam Soleimanzadeh, Rafael Wisniewski, and Stoyan Kanev. 2012. An optimization framework for load and power distribution in wind farms. *Journal of Wind Engineering and Industrial Aerodynamics* 107 (2012), 256–262.
- [34] Constantinos Sourkounis and Pavlos Tourou. 2013. Grid code requirements for wind power integration in Europe. In *Conference Papers in Energy*, Vol. 2013. Hindawi.
- [35] Vedrana Spudic, M Jelavic, M Baotic, and Nedjeljko Peric. 2010. Hierarchical wind farm control for power/load optimization. *The science of making torque from wind (TORQUE 2010)* (2010).
- [36] Ruben Stranders, Long Tran-Thanh, Francesco M. Delle Fave, Alex Rogers, and Nicholas R. Jennings. 2012. DCOPs and bandits: Exploration and exploitation in decentralised coordination. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. International Foundation for Autonomous Agents and Multiagent Systems, 289–296.
- [37] Siyu Tao, Andrés Feijóo, Jiemin Zhou, and Gang Zheng. 2020. Topology Design of an Offshore Wind Farm with Multiple Types of Wind Turbines in a Circular Layout. *Energies* 13, 3 (2020), 556.
- [38] William R. Thompson. 1933. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika* 25, 3/4 (1933), 285–294.
- [39] Mike T. van Dijk, Jan-Willem Wingerden, Turaj Ashuri, Yaoyu Li, and Mario A. Rotea. 2016. Yaw-Misalignment and its Impact on Wind Turbine Loads and Wind Farm Power Output. *Journal of Physics: Conference Series* 753, 6 (2016).
- [40] Timothy Verstraeten, Eugenio Bargiacchi, Pieter J. K. Libin, Jan Helsen, Diederik M. Roijers, and Ann Nowé. 2020. Multi-Agent Thompson Sampling for Bandit Applications with Sparse Neighbourhood Structures. *Scientific Reports* 10 (2020). <https://doi.org/10.1038/s41598-020-62939-3>
- [41] Timothy Verstraeten, Eugenio Bargiacchi, Pieter J. K. Libin, Diederik M. Roijers, and Ann Nowé. 2020. Thompson Sampling for Factored Multi-Agent Bandits. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. International Foundation for Autonomous Agents and Multiagent Systems, 2029–2031.
- [42] Timothy Verstraeten, Pieter J. K. Libin, and Nowé Ann. 2020. Fleet Control using Coregionalized Gaussian Process Policy Iteration. In *Proceedings of the 24rd European Conference on Artificial Intelligence (ECAI)*.
- [43] Timothy Verstraeten, Ann Nowé, Jonathan Keller, Yi Guo, Shuangwen Sheng, and Jan Helsen. 2019. Fleetwide data-enabled reliability improvement of wind turbines. *Renewable and Sustainable Energy Reviews* 109 (2019), 428–437.
- [44] JW Wagenaar, L Machiels, and J Schepers. 2012. Controlling wind in ECN’s scaled wind farm. *Proc. Europe Premier Wind Energy Event* (2012), 685–694.
- [45] World Wind Energy Association (WWEA). 2019. <https://library.wwindea.org/global-statistics/>.