On local minima in distributed energy scheduling

Astrid Nieße*, Jörg Bremer†, Sebastian Lehnhoff†

*R&D Division Energy, OFFIS – Institute for Information Technology, Escherweg 2, D-26121 Oldenburg, Germany
Email: {astrid.niesse}@offis.de
†Energy Informatics, Department of Computing Science, University of Oldenburg, D-26111 Oldenburg, Germany
Email: {joerg.bremer, sebastian.lehnhoff}@uni-oldenburg.de

Abstract—**Distributed energy scheduling constitutes a tough task for optimization algorithms, as the underlying problem structure is highdimensional, multimodal and non-linear. For this reason, metaheuristics and especially distributed algorithms have been in the focus of research for several years with promising results. The modeling of the distributed energy units' flexibility is a specific research task, with different concepts like comfort-level based approaches, enumeration of possible schedules, and continuous schedule representation using machine learning and decoder techniques. Although a continuous representation of flexibility has shown better results regarding the global optimization goal, there have been hints that the susceptibility to local minima traps enlarges compared to the enumeration of distinct schedules. In this contribution, we present an exemplary system for predictive scheduling of distributed energy units consisting of a continuous flexibility modelling approach and a fully distributed planning heuristic. A prestudy is presented, where we analyze the problem structure regarding local minima and describe planned work to reduce the heuristic's susceptibility to be kept in these.**

Index Terms—**Distributed Energy Scheduling, Agent-Based Control, Self-Organization, Unit Commitment, Local Minima, Meta-Heuristics.**

I. INTRODUCTION

IN FUTURE energy systems, distributed energy units like renewables, small combined heat and power (CHP) plants, and electrical storages are needed for both energy provision and grid stabilization purposes. Typically, these units are operated in an aggregated fashion by so-called virtual power plants (VPP), constituting the core concept for renewable energy systems [1]. One of the main challenges during operation of such a VPP arises from the complexity of the scheduling task due to the large amount and the diverse nature of energy units in the distribution grid [2].

The general optimization problem to be solved for this scheduling task is known as unit commitment problem [3] and constitutes a combinatorial optimization problem: Under given constraints specific for the respective energy units (local constraints) operation modes for each unit have to be chosen in such a way that the global optimization goal is reflected for the whole planning horizon. These kind of problems have been described as multiple-choice subset-sum problem [4]. It has been shown that for the given task of predictive scheduling the problem is weakly NP-complete, if the continuous operation of

energy units is discretized to operation states per time interval (schedules) [5]. Following this discretization, the optimality depends on the combination of the schedules of all energy units in each interval.

Thus, the problem of predictive scheduling in VPP shows the following characteristics: First, for day-ahead scheduling with a resolution of 96 time intervals the problem is highdimensional. Second, the combinations of schedules may show the same performance with different combinations of schedules – thus the problem is multimodal. Third, with the non-convex structure of the solution space that is given with the physical properties of the energy units, the problem has to be classified as non-linear. Due to this structure of the optimization problem, several heuristic optimization algorithms have been proposed and evaluated for predictive scheduling in VPP [5], [6], [7], [8].

With the concept of schedules as discretized operation modes, a reduced view on the potential flexibility of the respective energy unit is chosen: While some units may be operated in a fine-grained manner, the coarse grained solution space narrows schedule choices to a distinct set of schedules without tapping the units' full optimization potential. To this end, a support vector data description (SVDD) model and a decoder approach have been introduced [9], allowing for a continuous representation of flexibility combined with a targeted unconstrained search based on this surrogate model. Although the usage of the decoder largely helps to identify solutions that better use the units' flexibility [10], there has been evidence that the amount of local minima increases as well. With the general predisposition of heuristic algorithms to local minima [11] a more in-depth analysis of the relation of flexibility modelling and optimization performance is necessary: The combination of a search space representation and the heuristic using this search space has to be chosen carefully [12].

In this contribution we present a preliminary study on the predisposition of the combination of SVDD based flexibility modelling, decoder based search and the distributed optimization heuristic COHDA [5] for local minima and give an outline of planned work to analyze the dependencies in greater depth.

The rest of this position paper is structured as follows: In section II we give an introduction to different approaches to flexibility modelling, followed by an overview on distributed energy scheduling in section III. We then present a preliminary study on the occurance of local minima in different search space settings and conclude with an outline of planned work.

II. FLEXIBILITY MODELING

Flexibility modeling can be understood as the task of modeling constraints for energy units. Apart from global VPP constraints, constraints often appear within single energy components; affecting the local decision making. Popular methods treat constraints or aggregations of constraints as separate objectives or penalties, leading to a transformation into a (unconstrained) many-objective problem [13], [11].

For optimization approaches in smart grid scenarios, black-box models capable of abstracting from the intrinsic model have proved useful [14], [15]. The units do not need to be known at compile time. A powerful, yet flexible way of constraint-handling is the use of a decoder that gives a search algorithm hints on where to look for schedules satisfying local hard constraints [15], [16].

A decoder is a technique that gives algorithms hints on where to look for feasible solutions and thus allows for a targeted search. It imposes a relationship between a decoder solution and a feasible solution and gives instructions on how to construct a feasible solution [16]. Using directly a given set $\mathcal{X}$ of feasible schedules derived from a simulation model can already serve as a decoder [5] without a need for machine learning techniques to deduce a meta-model, a response surface, or similar. Each schedule is simply mapped on the most similar one from $\mathcal{X}$ with the lowest distance.

We regard a schedule of an energy unit as a vector $\boldsymbol{x} = (x_0, \ldots, x_d) \in \mathcal{F} \subset \mathbb{R}^d$ with each element x_i denoting mean power generated (or consumed) during the ith time interval.

A simple, yet easy way to find a surrogate model for the flexibility of an energy unit is to abstract from operation and constraints by maintaining a set $\mathcal{S}$ of feasible example schedules and use this set as abstraction layer. Any given schedule $\boldsymbol{x}$ is feasible iff $\boldsymbol{x} \in \mathcal{S}$. A relaxed variant might demand that $\exists \hat{\boldsymbol{x}} \in \mathcal{S} \bullet \|\hat{\boldsymbol{x}} - \boldsymbol{x}\| \leq \epsilon$ for some small, given threshold. We take the strict version.

So far, this surrogate is capable of checking feasibility when already given a schedule. In this way, the surrogate may tell apart feasible and infeasible schedules on behalf of the specific simulation model of the energy unit and thus already allows for an abstraction from any model specific implementation. On the other hand, it is not yet a sufficient constraint-handling technique as it still needs externally (e. g. by any optimization algorithm) generated schedules which can merely be checked. Hence, we need to guide an algorithm where to look for feasible schedules. A decoder can do this. For the set based approach, a decoder can simply be derived as follows:

$$\tau_{\text{set}} : \begin{cases} \mathbb{R}^d \to \mathcal{S} \\ \boldsymbol{x} \mapsto \boldsymbol{x}' \in \mathcal{S}, \|\boldsymbol{x} - \boldsymbol{x}'\| < \|\boldsymbol{x} - \boldsymbol{x}^*\| \ \forall \boldsymbol{x}^* \in \mathcal{S}, \boldsymbol{x}' \neq \boldsymbol{x}^* \end{cases}$$
(1)

In this way, any given (feasible or not) schedule $\boldsymbol{x}$ is mapped onto the nearest feasible schedule from $\mathcal{S}$. The granularity of this model depends on the cardinality of the set $\mathcal{S}$. The larger $|\mathcal{S}|$ the more the model resembles a continuously modeled feasible space. We will now discuss a way of modeling flexibility in a real continuous space.
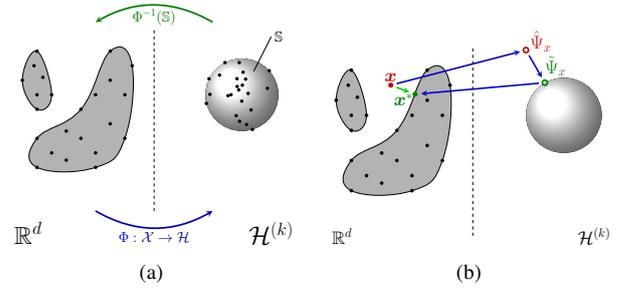
Fig. 1. General support vector decoder scheme for solution repair and constraint handling [19].

Fig. 1 shows the idea of a support vector decoder starting with a set of feasible example schedules derived from a simulation model of the respective energy unit and using it as a stencil for the region that contains just feasible schedules.

A training set $\mathcal{X}$ containing only valid schedules, can e. g. be derived after a sampling approach from [17]. From such training set, a SVDD [18] derives a geometrical description of the sub-space that contains the given data (in our case: the set of feasible schedules). Given a set of data samples, the enclosing envelope can be derived as follows: After mapping the data to a high dimensional feature space, the smallest enclosing ball in this feature space is determined. When mapping back the ball to data space, it forms a set of contours enclosing the given data sample.

This task is achieved by determining a mapping $\Phi : \mathcal{X} \subset \mathcal{F} \subset \mathbb{R}^d \to \mathcal{H}; \ x \mapsto \Phi(x)$ such that all data from a sample $\mathcal{X}$ is mapped to a minimal hypersphere in $\mathcal{H}$. The minimal sphere with radius R and center a in $\mathcal{H}$ that encloses $\{\Phi(\boldsymbol{x}_i)\}_N$ can be derived from minimizing $\|\Phi(\boldsymbol{x}_i) - a\|^2 \leq R^2 + \xi_i$ with slack variables $\xi_i \geq 0$ for a smoother ball.

After some relaxations one gets two main outcomes: the center $a = \sum_i \beta_i \Phi(\boldsymbol{x}_i)$ (with $\boldsymbol{\beta}$ weighting the impact of different schedules) of the minimal sphere in terms of an expansion into $\mathcal{H}$ and a function that allows to determine the distance of the image of an arbitrary point from $a \in \mathcal{H}$, calculated in $\mathbb{R}^d$ is derived: $R^2(\boldsymbol{x}) = 1 - 2 \sum_i \beta_i k_G(\boldsymbol{x}_i, \boldsymbol{x}) + \sum_{i,j} \beta_i \beta_j k_G(\boldsymbol{x}_i, \boldsymbol{x}_j)$. Because all support vectors are mapped onto the surface of the sphere, the sphere radius $R_{\mathbb{S}}$ can be easily determined by the distance of an arbitrary support vector to the center a. Thus the feasible region can now be modeled as $\mathcal{F} = \{\boldsymbol{x} \in \mathbb{R}^d | R(\boldsymbol{x}) \leq R_{\mathbb{S}}\} \approx \mathcal{X}$.

The model can be used as a black-box that abstracts from any explicitly given form of constraints and allows for a decision on whether a given solution is feasible or not. From the support vector model, a decoder can be derived automatically. The set of feasible schedules is represented as pre-image of a high-dimensional ball $\mathbb{S}$. Fig. 1(a) shows the geometric situation. This representation has some advantageous properties. Although the pre-image might be some arbitrary shaped non-continuous blob in $\mathbb{R}^d$, the high-dimensional representation is a ball and thus geometrically easier to handle.

If a schedule is feasible it is inside the feasible region (grey area on the left in Fig. 1(b)). Thus, the schedule is
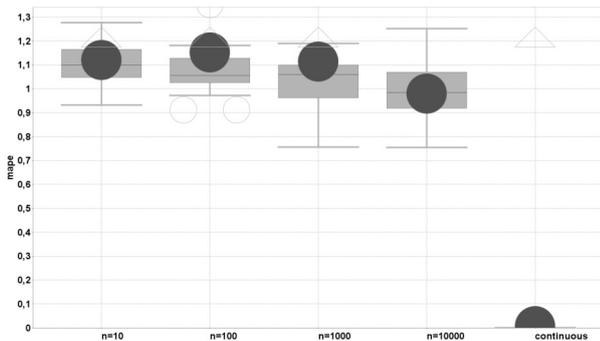
Fig. 2. Solution quality (as mean absolute percentage error, mape) as a function of the granularity of the flexibility model.

inside the pre-image (representing the feasible region) of the ball and thus its high-dimensional image lies inside the ball. An infeasible schedule (e. g. $\boldsymbol{x}$ in Fig. 1(b)) lies outside the feasible region and thus its image $\hat{\Psi}_x$ lies outside the ball. But we know some relations: the center of the ball, the distance of the image from the center and the radius of the ball. Hence, we can move the image of an infeasible schedule along the difference vector towards the center until it touches the ball. Finally, one can calculate the pre-image of the moved image $\tilde{\Psi}_x$ and get a schedule at the boundary of the feasible region: a repaired schedule $\boldsymbol{x}^*$ that is now feasible. No mathematical description of the original feasible region or of the constraints is needed to do this. More sophisticated variants of transformation are e. g. given in [9]. Formally, the decoder τ is given by

$$\tau : \mathbb{R}^d \to \mathcal{F}_{U_i} \subseteq \mathbb{R}^d, \ \boldsymbol{x} \mapsto \tau(\boldsymbol{x}) = \hat{\Phi}^{-1} \circ \tau_{\text{move}} \circ \hat{\Phi}. \quad (2)$$

Using any decoder, the global scheduling problem can be transformed into a formulation that is unconstrained regarding local constraints. Apart from finding a combination of schedules whose sum resembles a given target power profile best, further objectives are usually integrated due to the many-objective nature of energy scheduling. In the following, we consider predictive scheduling, where the goal is to select exactly one schedule $\boldsymbol{x}_i$ for each energy unit U_i from a search space of feasible schedules with respect to a future planning horizon, such that a global objective (e. g. resembling a target power profile) is optimized by the sum of individual contributions [10]. A basic formulation of the scheduling problem is given by

$$\delta \left(\sum_{i=1}^{m} \boldsymbol{x}_i, \boldsymbol{\zeta} \right) \to \min; \text{ s.t. } \boldsymbol{x}_i \in \mathcal{F}^{(U_i)} \ \forall U_i \in \mathcal{U}. \quad (3)$$

In equation (3) δ denotes a distance measure for evaluating the difference between the aggregated schedule of the VPP and the desired target schedule $\boldsymbol{\zeta}$. To compare results and for scalability reasons we used the mean absolute percentage error (mape) $\delta(\boldsymbol{x}, \boldsymbol{\zeta}) = \frac{100}{d} \sum_{i=1}^{d} \left| \frac{\zeta_i - x_i}{\zeta_i} \right|$.

To each energy unit U_i exactly one schedule $\boldsymbol{x}_i$ has to be assigned. $\mathcal{F}^{(U_i)}$ denotes the individual set of feasible schedules

that are operable for unit U_i without violating any (technical) constraint. Solving this problem without unit independent constraint handling leads to specific implementations that are not suitable for handling changes in VPP composition or unit setup without having changes in the implementation of the scheduling algorithm [15]. Using a decoder for constraint handling one can now rephrase the optimization problem as

$$\delta \left(\sum_{i=1}^{m} \tau_i(\boldsymbol{x}_i), \boldsymbol{\zeta} \right) \to \min, \quad (4)$$

where τ_i is the decoder function of unit i that produces feasible schedules from $\boldsymbol{x} \in [0, x_{max}]^d$ resulting in schedules that are operable by that unit. Please note, that this is a constraint free formulation. With this problem formulation, many standard algorithms for optimization can be easily adapted as there are no constraints (apart from a simple box constraint $\boldsymbol{x} \in [0, x_{max}]^d$) to be handled and no domain specific implementation (regarding the energy units and their operation schedules) has to be integrated in the optimization algorithm. Equation (4) is used as a surrogate objective to find the solution to the constrained optimization problem Eq. (3).

Fig. 2 shows some mean optimization results obtained by using different flexibility models. Four discrete flexibility models (cf. decoder Eq. 1) have been used with sets ranging from $n = 10$ to 10000 elements (schedules). Additionally one fully continuous flexibility model has been used. It is obvious that the optimization result gets better the more choices are contained in the model. Moreover, a significant improvement can be made by transition to a fully continuous model. On the other hand, with growing granularity, the variance of the results and the number of worse outliers also grow indicating a growing risk to premature convergence. Also in the continuous case outliers indicating results worse than with the coarse grained model (upper triangle) are present. This result also reproduces preliminary results from [10].

It seems immediately advantegeous to use continuous or at least fine grained models to obtain better results. On the other hand, this design decision entails a growing complexity and a larger modality to the objective making the problem harder to solve due to problems with premature convergence. We will show this in greater detail in the following section.

III. DISTRIBUTED ENERGY SCHEDULING

Heuristic and especially distributed and agent-based algorithms have been in the focus of research within the last years. Nevertheless, the notion of what constituted a distributed algorithms still differs a lot: Whereas early publications focussed on the atomic units (mainly named agents) as a gateway to distributed energy units, many hierarchical approaches have been presented for both energy market aspects [20], [21], [22] as balancing approaches [23]. In recent years, some fully distributed algorihms have been presented, with the individual agents being capable of solution generation and evaluation without a centralized components [24], [25].

The Combinatorial Optimization Heuristic for Distributed Agents (COHDA, originally introduced in [5]) has been evaluated for predictive scheduling problems in VPPs and has been chosen for our analysis of local minima susceptance with a continuous flexibility model. Although parts of the planned work regarding local minima convergence reduction is specific for COHDA (the algorithmic parts), the general problem analysis is valid for all heuristic distributed algorithms using a continuous flexibility model. In the following, COHDA is described in a concise fashion, based on the description in [26].

A. Introducing COHDA

The key concept of COHDA is an asynchronous iterative approximate best-response behavior, where each agent – representing a decentralized energy unit – reacts to updated information from other agents by adapting its own selected schedule with respect to the global objective function (OF). All agents $a_i \in A$ initially only know their own respective set of schedules S_i. From an algorithmic point of view, the difficulty of the problem is given by the distributed nature of the system in contrast to the task of finding a common allocation of schedules for a global target power profile.

Thus, the agents coordinate by updating and exchanging information about each other. For privacy and communication overhead reasons, the potential flexibility (i.e. the set of feasible schedules) S_i is not communicated as a whole by an agent a_i. Instead, the agents communicate single selected schedules within the approach as described in the following.

First of all, the agents are placed in an artificial communication topology based on the small-world scheme, such that each agent is connected to a non-empty subset of other agents. This overlay topology might be a ring in the least connected variant.

Each agent a_i collects two distinct sets of information: on the one hand the believed current configuration γ_i of the system (that is, the most up to date information a_i has about currently selected schedules of all agents), and on the other hand the best known combination γ_i^* of schedules with respect to the global objective function it has encountered so far.

Beginning with an arbitrarily chosen agent by passing it a message containing only the global objective (i. e. the target power profile), each agent repeatedly executes the three steps *perceive*, *decide*, *act* (cf. [26]) as visualized in Fig. 3:

1) **perceive:** When an agent a_i receives a message κ_p from one of its neighbors (say, a_p), it imports the contents of this message into its own memory.
2) **decide:** The agent then searches S_i for the best schedule regarding the updated system state γ_i and the global objective function. Local constraints are taken into account in advance. Details regarding this procedure have been presented in [19]. If a schedule can be found that satisfies both the objectives, a new schedule selection is created. If the resulting modified system state γ_i yields a better rating than the current solution candidate γ_i^*, a new solution candidate is created based on γ_i. Otherwise
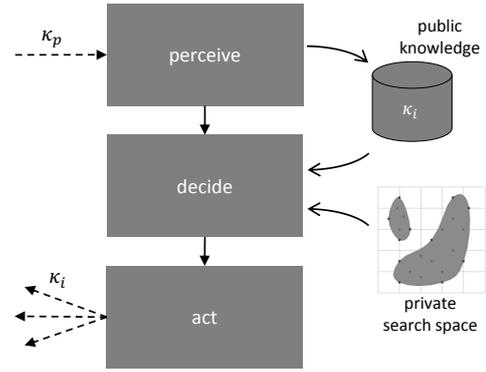
Fig. 3. The *perceive–decide–act* behavioral pattern in COHDA from the point of view of an agent a_i [19].

the old solution candidate still reflects the best schedule combination regarding the global objective the agent is aware of, so the just created schedule selection is discarded and the agent reverts to its schedule selection stored in γ_i^*.
3) **act:** If γ_i or γ_i^* has been modified in one of the previous steps, the agent finally broadcasts these to its neighbors in the communication topology.

Following this behavior, only small subsets of the sets of feasible schedules S_i are communicated by the agents. During this process, for each agent a_i, its observed system configuration γ_i as well as solution candidate γ_i^* are filled successively. After producing some intermediate solutions, the heuristic eventually terminates in a state where for all agents γ_i as well as γ_i^* are identical, and no more messages are produced by the agents. At this point, γ_i^* is the final solution of the heuristic and contains exactly one schedule selection for each agent.

B. Convergence of COHDA in local minima

During the above described phases *perceive*, *decide*, and *act*, a better solution candidate always prevails over inferior solution candidates, and the system always terminates with convergence. Due to the asynchronous search in the solution space, inferior local optima are discarded as long as other agents are able to identify better solutions.

The susceptibility to local minima convergence depends on both the structure of the solution space and the search algorithm. With a course-grained representation of schedules as enumerated set, COHDA always reached near-optimal convergence [5]. In the following we present results of two application examples of COHDA where this behavior can not be guaranteed and thus motivate a deeper analysis of local minima susceptibility.

1) Continuous scheduling: In [19], COHDA has been coupled with a continuous flexibility modeling approach, thus allowing for a targeted search within that search space and delivering a fine-grained flexibility description. This system has been applied for the task of continuous scheduling in VPP: after an initial predictive scheduling, incidents (DER

breakdowns, prognoses faults, ...) may render the initial schedules infeasible. In this case, the infeasible schedules have been marked as invalid and an adapted variant of COHDA has been restarted for a new cooperative search.

In Fig. 4 an exemplary outcome of such a continuous scheduling process with 10 CHP outages is shown. On the x-axis, the solution evaluation is chronologically ordered: Each time an agents evaluates a VPP schedule within the OF, this is logged within the system. On the y-axis, the expected product delivery performance is given. A value of 1.0 means perfect product delivery (i.e. 150 kWh in all 4 product hours). In the left part of the diagram, the initial planning process is shown: The agents yield an expected product delivery performace of 0.99 in the day ahead planning process. For all following values, the simulated time is given on top of the diagram for ease of understanding. Incidents are depicted using arrows.
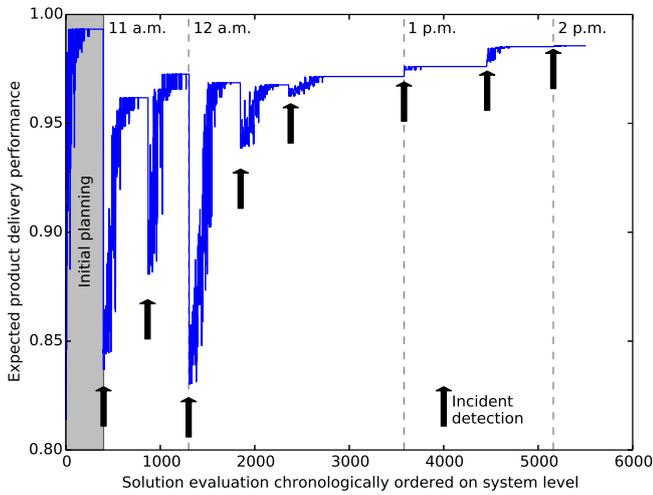
Fig. 4. Example for incident detection and expected product delivery performance [27].

Fig. 5. Product fulfillment, simulation duration and cooperation overhead with rising sensitivity values, adapted from [28]. Each simulation set shows the result of 100 simulation runs as boxplot.

In the first half of product delivery, each incident leads to an initially reduced product delivery performance. The agents manage to enhance this value within the cooperative search for a new VPP schedule though. Convergence of the processes can be recognized from the plateaus before the next incident.

During the second half of product delivery, an effect not yet explained is observed: with an incident, the product delivery performance does not decrease as expected, but increases to a better value at once. A possible explanation for this might be that the incidents lead to a restart of the heuristic and thus an escape from a local minimum. In Fig. 4 only one exemplary optimization restart process is given. During the examination of the approach presented in [27] the effect of a better performance after disturbance has been observed several times, even with the effect of a better performance after an incident compared to the initial predictive scheduling.

2) Reducing evaluation sensitivity: Due to the continuous representation of the flexibility, it is not possible to calculate the global optimum even in small scenarios. Therefore, the analysis if a simulation has converged within a local minimum cannot be done using a brute force approach in the combined system of COHDA and SVDD based flexibility description. To analyze the susception to local minima of this combined system, a trick can be chosen: Inspite of analysing the problem structure, the sensitivity of the agents when analysing the problem structure and searching for a new solution candidate can be changed by artificially reducing the precision within the OF.

In Fig. 5, results of this approach are shown for a scenario with 100 energy units (CHP and PV plants). The agents had to plan a product of 600 kWh to be delivered within one hour. The results for product fulfillment (with a value of 1.0 denoting perfect fulfillment of 600 kWh and values above 1.0 denoting overfulfillment), simulation duration until convergence and cooperation overhead (total amount of messages until convergence). For all settings of the sensitivity value, the problem structure has been the same. The precision of the optimization evaluation has been modified from 0.0 (i.e. full precision) to 0.01. It can be seen that with a larger sensitivity value, both simulation duration and cooperation overhead go down. The result quality though, given as product fulfillment, seems to be instable: While with a sensitivity value of 0.005 all simulation runs lead to a product fulfillment close to 1.0, the opposite is the case with a value of 0.01: All simulation runs converge with a product fulfillment of 1.2, a local minimum that already has been encountered in the other runs. Obviously, the system has been trapped in a local minimum in the last simulation series.

Up to now, we did not analyze the effects of combining the continuous flexibility representation with COHDA in sufficient detail. With these results, it has been decided to analyze the context of continuous flexibility modeling, COHDA and local minima susceptibility on more detail.
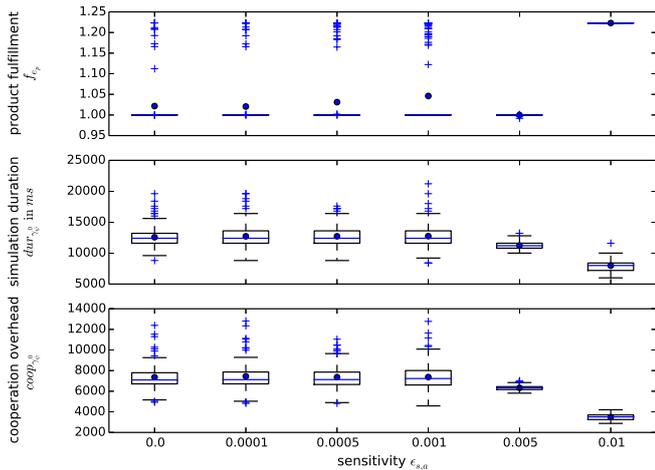
IV. PRESTUDY: OCCURENCE OF LOCAL MINIMA

In order to underpin our conjectures regarding the increase of local minima with introducing fine grained flexibility modeling, we investigated the changes in problem complexity with means from fitness landscape analysis.

A. Structure analysis

We start by analyzing the structure of fitness landscapes. A landscape $\mathcal{L}$ is defined by the search space $\mathcal{S}$, a neighborhood relation $\mathcal{N}$ and an error function (equivalently: the fitness) ν. Plotting ν over all $x \in \mathcal{S}$ yields the landscape that is scanned by some optimization algorithm for the optimal point. The structure of this landscape determines (premature) convergence and thus whether the algorithm succeeds with a sufficiently small budget or gets stuck within some local optimum.

One can analyze the structure of this landscape e. g. by scrutinizing the error correlation of neighboring candidate solutions [29]. Neighboring solutions from flat regions of the landscape exhibit a higher correlation than solutions from rugged parts of the landscape. Thus, the correlation can be seen as some measure for the ruggedness of $\mathcal{L}$.

We scrutinize the autocorrelation of random paths on the landscape (random walk through $\mathcal{N}$) [30]. Let $X = (x_{ij})$ be a solution of the predictive scheduling problem Eq. (4) with each row denoting a schedule for unit i and each element representing mean active power during the respective time interval j. Let $\{f_t\}_{t=1}^n$ be a sequence of n objective values, sampled as follows: starting from a randomly chosen solution $X_1 \in \mathcal{S}$ successive, neighboring solutions X_{t+1} are generated after [31] by altering each element in X_t by adding or subtracting 0.1 with a probability of $1/3$ each. It is ensured that the solution stays within the feasible region and that at least one element of X_t is altered. The series $F = \{f_t\}_{t=1}^n$ now contains values $f_t = \nu(X_t)$, with $\nu(X) = \delta(\sum_{x_i \in X} x, z)$.

Now one can calculate the autocorrelation

$$\rho(\sigma) = \frac{E[f_t f_{t+\sigma}] - E[f_t]E[f_{t+\sigma}]}{V[f_t]} \tag{5}$$

for a given path length σ, with $E[f_t]$ and $V[f_t]$ denoting expectation and variance respectively. Moreover, [30] defines the correlation length $\lambda = -\frac{1}{ln(\rho(1))}$, denoting the mean distance (in the sense of neighboring hops of $\mathcal{N}$) from which the majority of the solutions is no longer correlated [29]. The correlation length can also be interpreted as the expected maximum width of flat valleys in the landscape.

B. Information analysis

Analyzing the correlation of random paths on the landscape yields an impression of the structure. In [29] an extended analysis is proposed based on entropy measures on $\{f_t\}_{t=1}^n$. Founded on ideas from algorithmic information theory [32] and the entropy from Shannon [33], a characterization of the distribution and number of local optima along the path is given as a measure of the complexity of the fitness landscape.

For the following indicators, random paths on the landscape are seen as ensemble of base elements. Three element types

(token) can be distinguished: flat areas (neighboring points have similar fitness), isolated points (surrounded merely by better, or worse points), and slope points (neither isolated nor flat). In a first step, each path is transformed in a sequence of tokens $S(\epsilon) = s_1 s_2 \ldots s_n$ over the alphabet $\{\bar{1}, 0, 1\}$ by

$$S_i = \psi_{f_t}(i, \epsilon) = \begin{cases} \bar{1}, & \text{if } f_i - f_{i-1} < -\epsilon \\ 0, & \text{if } |f_i - f_{i-1}| \leq \epsilon \\ 1, & \text{if } f_i - f_{i-1} > \epsilon \end{cases} \tag{6}$$

for a given $\epsilon \in [0, \max f_t]$ (cf. [29]) A string $S(\epsilon)$ then contains information on the structure of the landscape along a randomly chosen path. Now one can define an object by two successive tokens in the string. For example, the sequence $\bar{1}1$ denotes a change from downslope to upslope and thus a trough. The entropy measure for such an ensemble of objects can be determined after [29]:

$$H(\epsilon) = -\sum_{p \neq q} P_{[pq]} \log_6 P_{[pq]}, \tag{7}$$

with $P_{[pq]}$ denoting the frequency of the occurrence of the sequence pq in $S(\epsilon)$. The modality of the objective function can also be derived from $S(\epsilon)$. As opposed to the entropy which is a measure of the diversity of objects along the path, the modality must be measured by a classification of objects in order to determine the number of (local) optima. First, the partial information content is determined [29]. To achieve this, the string $S(\epsilon)$ is transformed into $S'(\epsilon) = o_1 o_2 \ldots o_\mu$ over the alphabet $\{\bar{1}, 1\}$. This yields the shortest string that represents the alternations from uphill to downhill changes along the path. The partial information content is recursively defined by [29]:

$$M(\epsilon) = \frac{\mu}{n} = \frac{|S'(\epsilon)|}{|S(\epsilon)|} \in [0, 1]. \tag{8}$$

A value of 1 denotes the maximum modality. The absolute number of (local) optima (according to a given ϵ) can be derived by $\lfloor (n \cdot M(\epsilon))^{-2} \rfloor$. All these measures are sensitive to the choice of ϵ. Small values lead to a higher sensitivity to changes in fitness between neighbouring solution. The smallest value of ϵ that lets all differences vanish is called information stability [29] (fully flat error function).

C. Results

As a unit model we used a co-generation model that has already served in several studies and projects for evaluation [34], [35], [10], [36]. This model comprises a micro CHP bundled with a thermal buffer store. Constraints restrict power band, buffer charging, gradients, min. on and off times, and satisfaction of thermal demand. Thermal demand is determined by simulating a detached house. For each agent the model is individually (randomly) configured with state of charge, weather condition, temperature range, allowed operation gradients, and similar. We experimented with problem sizes of 20 energy units for scheduling.

First, we scrutinized the complexity of the fitness landscape using the correlation analysis. Figure 6(a) shows the result.
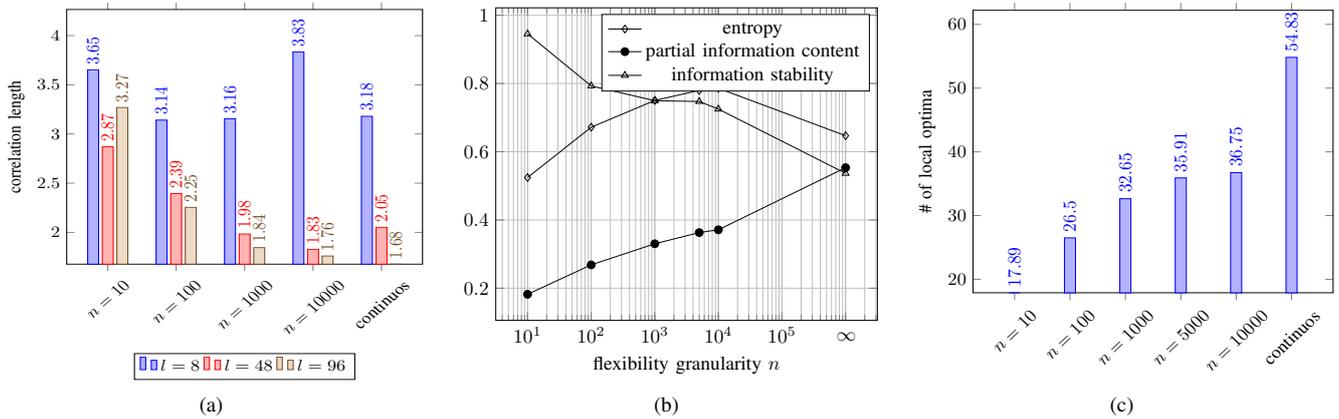
Fig. 6. Experimental results of the fitness landscape analysis.

TABLE I
COMPARISON OF DIFFERENT TYPES OF ENERGY UNITS.

unit type	entropy	part. inf.	# local optima	inf. stability
heat pump	0.708 ± 0.045	0.495 ± 0.039	49.04 ± 3.91	0.462 ± 0.095
boiler	0.69 ± 0.046	0.459 ± 0.039	45.48 ± 3.93	0.372 ± 0.171
cool storage	0.501 ± 0.096	0.171 ± 0.050	16.74 ± 5.03	0.709 ± 0.173
chp	0.354 ± 0.097	0.113 ± 0.044	11.13 ± 4.42	0.943 ± 0.104

Comparing the correlation length of the scheduling problem with different granularity of flexibility modeling, one can derive that the complexity grows with finer grained models. The correlation length almost vanishes with the continuous model. Thus, neighboring solutions are mostly uncorrelated with fine grained models. This result is immediately apparent as the coarse grained models map many neighboring solutions onto the very same phenotype solution, resulting in a higher correlation. The experiment has been conducted for different schedule lengths. The effect is different for different problem complexities (cf. Fig. 6(a)), because shorter schedules exhibit less regions with high information content [15].

Fig. 6(b) shows results of the information analysis. Whereas information stability and partial information indicate a rising modality with rising granularity in flexibility modeling, the entropy slightly recedes for the continuous case. This might be founded in vanishing steps that are prominent in the discrete case. The result in Fig. 6(c) most prominently shows the difference in complexity when comparing the discrete and the continuous flexibility models: the mean (absolute) number of local minima that are encountered along a random path across the fitness landscape. Although the absolute number depends on path length (200 in this case) and on the choice of ϵ, the general trend is still always the same for this experiment.

Obviously, also the types of modeled energy units have an impact on these results. Table I compares the results for different problems with different types of energy units.

V. CONCLUSION AND FURTHER WORK

Predictive scheduling with its combinatorial optimization problem is a frequent task in future virtual power plants. As

any algorithm implementation should be designed independently from the aggregated energy units an abstraction layer renders indispensable. Flexibility modeling with decoders allows integrating any energy unit into the optimization model without domain specific knowledge on the possible future operations, technical constraints or on the structure of the individual search space of feasible schedules by building up a constraint-free search space.

Flexibility modeling may be achieved with different surrogate techniques. First we presented hints from earlier studies to problems regarding local minima susceptance with a fine-grained decoder based flexibility model. Then we scrutinized the effect of two possible implementations and showed the dependencies on the granularity of the model: Fine grained or continuous models allow for higher precision and better optimization results and are thus desirable. On the other hand, such fine grained models introduce complexity by additional local minima in the objective.

These results lead up to further work, with the general goal of (a) getting a deeper understanding of the predisposition of fine-grained flexibility modeling combined with distributed algorithms for predictive scheduling to premature local minima convergence and (b) analzing algorithmic adaptations to COHDA to reduce premature local minima convergence rates. The following studies will be conducted with these goals:

1) Topology adaptations: For COHDA, different overlay topologies have been analyzed in detail. After a first setup of the agents' neighborhood though, no adaptations have been made during an ongoing optimization. In a simulation study we want to analyze the effect of dynamically adapting the communication topology on local minima convergence.

2) Artificial restart: There have been indications that the result quality benefits from a restart of the cooperative search within COHDA (see figure 4). To fully understand this effect, we plan to artificially restart an already terminated optimization (e. g. by schedule invalidation, minor target adaptation, or similar).

3) Sensitivity adaptations: In preliminary studies we could

show that introducing a sensitivity factor to the OF leads to changes within the local minima convergence. In these studies though, the factor has been chosen statically. In future work, we plan to dynamically adapt this factor, using approaches like simulated annealing.

Later work could include a change in the reaction delay of the agents and a temporary acceptance of deprecated solutions. For the latter, the convergence has to be assured.

REFERENCES

[1] H.-J. Appelrath, H. Kagermann, and C. Mayer, Eds., *Future Energy Grid. (acatech STUDY).* acatech, Munich, 2012.

[2] S. McArthur, E. Davidson, V. Catterson, A. Dimeas, N. Hatziargyriou, F. Ponci, and T. Funabashi, "Multi-agent systems for power engineering applications – Part I: Concepts, approaches, and technical challenges," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 1743–1752, 2007. doi: 10.1109/TPWRS.2007.908471

[3] N. Padhy, "Unit Commitment - A Bibliographical Survey," *IEEE Transactions on Power Systems*, vol. 19, no. 2, pp. 1196–1205, May 2004. doi: 10.1109/TPWRS.2003.821611

[4] D. Pisinger, "A minimal algorithm for the multiple-choice knapsack problem," *European Journal of Operational Research*, vol. 83, pp. 394–410, 1994.

[5] C. Hinrichs, S. Lehnhoff, and M. Sonnenschein, "A Decentralized Heuristic for Multiple-Choice Combinatorial Optimization Problems," in *Operations Research Proceedings 2012.* Springer, 2014. doi: 10.1007/978-3-319-00795-3_43. ISBN 978-3-319-00795-3 pp. 297–302.

[6] A. Feliachi and R. Belkacemi, "Intelligent Multi-agent system for Smart Grid Power Management," in *Smart Power Grids 2011*, A. Keyhani and M. Marwali, Eds. Springer, 2012, pp. 515–542.

[7] P. Vrba, V. Mařík, P. Siano, P. Leitão, G. Zhabelova, V. Vyatkin, and T. Strasser, "A Review of Agent and Service-Oriented Concepts Applied to Intelligent Energy Systems," *IEEE Transactions of Industrial Informatics*, vol. 10, no. 3, pp. 1890–1903, 2014.

[8] G. Anders, F. Siefert, and W. Reif, "A System of Systems Approach to the Evolutionary Transformation of Power Management Systems," in *43. Jahrestagung der Gesellschaft für Informatik e.V. (GI), Informatik angepasst an Mensch, Organisation und Umwelt.* Koblenz: Gesellschaft für Informatik, Köllen Druck+Verlag, 2013.

[9] J. Bremer and M. Sonnenschein, "Constraint-handling for optimization with support vector surrogate models – a novel decoder approach," in *Proceedings of the 5th International Conference on Agents and Artificial Intelligence (ICAART)*, J. Filipe and A. Fred, Eds., vol. 2. Barcelona, Spain: SciTePress, 2013. doi: 10.5220/0004241100910100 pp. 91–105.

[10] C. Hinrichs, J. Bremer, and M. Sonnenschein, "Distributed Hybrid Constraint Handling in Large Scale Virtual Power Plants," in *IEEE PES Conference on Innovative Smart Grid Technologies Europe (ISGT Europe 2013).* IEEE Power & Energy Society, 2013.

[11] A. Smith and D. Coit, *Handbook of Evolutionary Computation.* Department of Industrial Engineering, University of Pittsburgh, USA: Oxford University Press and IOP Publishing, 1997, ch. Penalty Functions, p. Section C5.2.

[12] J. Brownlee, *Clever Algorithms*, 1st ed. LuLu, 2011. ISBN 9781446785065

[13] O. Kramer, "A review of constraint-handling techniques for evolution strategies," *Appl. Comp. Intell. Soft Comput.*, vol. 2010, pp. 1–19, January 2010. doi: 10.1155/2010/185063

[14] F. Gieseke and O. Kramer, "Towards non-linear constraint estimation for expensive optimization," in *Applications of Evolutionary Computation*, ser. Lecture Notes in Computer Science, A. Esparcia-Alcázar, Ed. Springer Berlin Heidelberg, 2013, vol. 7835, pp. 459–468.

[15] J. Bremer and M. Sonnenschein, "Model-based integration of constrained search spaces into distributed planning of active power provision." *Comput. Sci. Inf. Syst.*, vol. 10, no. 4, pp. 1823–1854, 2013.

[16] C. A. Coello Coello, "Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: a survey of the state of the art," *Computer Methods in Applied Mechanics and Engineering*, vol. 191, no. 11-12, pp. 1245–1287, Jan. 2002. doi: 10.1016/S0045-7825(01)00323-1

[17] J. Bremer and M. Sonnenschein, "Sampling the search space of energy resources for self-organized, agent-based planning of active power provision," in *27th International Conference on Environmental Informatics, EnviroInfo 2013*, B. Page, A. G. Fleischer, J. Göbel, and V. Wohlgemuth, Eds. Hamburg, Germany: Shaker, 2013, pp. 214–222.

[18] D. M. J. Tax and R. P. W. Duin, "Support vector data description," *Mach. Learn.*, vol. 54, no. 1, pp. 45–66, 2004. doi: http://dx.doi.org/10.1023/B:MACH.0000008084.60811.49

[19] A. Nieße, J. Bremer, C. Hinrichs, and M. Sonnenschein, "Local Soft Constraints in Distributed Energy Scheduling," in *Proceedings of the 2016 Federated Conference on Computer Science and Information Systems (FEDCSIS).* IEEE, 2016. doi: 10.15439/2016F76 pp. 1517–1525.

[20] G. Anders, A. Schiendorfer, F. Siefert, J.-P. Steghöfer, and W. Reif, "Cooperative Resource Allocation in Open Systems of Systems," *ACM Transactions on Autonomous and Adaptive Systems*, vol. 10, no. 2, pp. 1–44, 2015. doi: 10.1145/2700323 ACM

[21] M. Tröschel, *Aktive Einsatzplanung in holonischen Virtuellen Kraftwerken.* Oldenburg: OlWIR, Oldenburger Verl. für Wirtschaft, Informatik und Recht, 2010. ISBN 978-3-939704-55-3

[22] P. Papadopoulos, N. Jenkins, L. M. Cipcigan, I. Grau, and E. Zabala, "Coordination of the Charging of Electric Vehicles Using a Multi-Agent System," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 1802–1809, dec 2013. doi: 10.1109/TSG.2013.2274391

[23] S. Lehnhoff, *Dezentrales vernetztes Energiemanagement - Ein Ansatz auf Basis eines verteilten Realzeit-Multiagentensystems.* Vieweg + Teubner, 2010.

[24] E. Pournaras, "Multi-level Reconfigurable Self-organization in Overlay Services," Ph.D. dissertation, TU Delft. ISBN 9789461860989 2013.

[25] B. Schäfer, M. Matthiae, M. Timme, and D. Witthaut, "Decentral Smart Grid Control," *New Journal of Physics*, no. JANUARY, 2015. doi: 10.1088/1367-2630/17/1/015002

[26] A. Nieße, S. Beer, J. Bremer, C. Hinrichs, O. Lünsdorf, and M. Sonnenschein, "Conjoint Dynamic Aggregation and Scheduling Methods for Dynamic Virtual Power Plants," in *Proceedings of the 2014 Federated Conference on Computer Science and Information Systems*, ser. Annals of Computer Science and Information Systems, M. Ganzha, L. A. Maciaszek, and M. Paprzycki, Eds., vol. 2. IEEE, 2014. doi: 10.15439/2014F76. ISBN 978-83-60810-58-3 pp. 1505–1514.

[27] A. Nieße and M. Sonnenschein, "A Fully Distributed Continuous Planning Approach for Decentralized Energy Units," in *45. Jahrestagung der Gesellschaft für Informatik e.V. (GI), Informatik, Energie und Umwelt.* Cottbus: Gesellschaft für Informatik, Köllen Druck+Verlag, 2015.

[28] A. Nieße, "Verteilte kontinuierliche Einsatzplanung in dynamischen virtuellen Kraftwerken," Ph.D. dissertation, Carl von Ossietzky Universität Oldenburg, Oldenburg, 2015.

[29] V. K. Vassilev, T. C. Fogarty, and J. F. Miller, "Information characteristics and the structure of landscapes," *Evol. Comput.*, vol. 8, no. 1, pp. 31–60, Mar. 2000. doi: 10.1162/106365600568095

[30] E. Weinberger, "Correlated and uncorrelated fitness landscapes and how to tell the difference," *Biological Cybernetics*, vol. 63, no. 5, pp. 325–336, 1990. doi: 10.1007/BF00202749

[31] G. Merkuryeva and V. Bolshakovs, "Benchmark fitness landscape analysis," *International Journal of Simulation Systems, Science & Technology (IJSSST)*, vol. 12, no. 2, pp. 38–45, 2011.

[32] G. J. Chaitin, *Algorithmic information theory*, ser. Cambridge tracts in theoretical computer science. Cambridge, Cambridgeshire, New York: Cambridge University Press, 1987. ISBN 0-521-34306-2

[33] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, 1948.

[34] J. Bremer, B. Rapp, and M. Sonnenschein, "Encoding distributed Search Spaces for Virtual Power Plants," in *IEEE Symposium Series in Computational Intelligence 2011 (SSCI 2011)*, Paris, France, 4 2011. doi: 10.1109/CIASG.2011.5953329

[35] J. Neugebauer, O. Kramer, and M. Sonnenschein, "Classification cascades of overlapping feature ensembles for energy time series data," in *Proceedings of the 3rd International Workshop on Data Analytics for Renewable Energy Integration (DARE'15), ECML/ PKDD 2015.* Springer, 2015.

[36] A. Nieße and M. Sonnenschein, "Using grid related cluster schedule resemblance for energy rescheduling - goals and concepts for rescheduling of clusters in decentralized energy systems." in *SMARTGREENS*, B. Donnellan, J. F. Martins, M. Helfert, and K.-H. Krempels, Eds. SciTePress, 2013. ISBN 978-989-8565-55-6 pp. 22–31.