Methoden Open Access

Simon Wenzel*, Radoslav Paulen, Goran Stojanovski, Stefan Krämer, Benedikt Beisheim, and Sebastian Engell

Optimal resource allocation in industrial complexes by distributed optimization and dynamic pricing

Optimale Ressourcenallokation in industriellen Prozessen durch verteilte Optimierung und dynamische Preise

DOI 10.1515/auto-2016-0003
Received January 15, 2016; accepted April 27, 2016

Abstract: We address the distributed hierarchical optimization of industrial production complexes where the individual plants exchange resources via networks. Due to the site-wide couplings a centralized or a distributed hierarchical optimization is needed to achieve the best overall performance of the site and to balance the networks of the shared resources. We discuss market-like algorithms that set prices of the shared resources in order to influence the individual optimizers so that the overall operation converges to the site-wide optimum. A novel algorithm for price adjustment based on the quadratic approximation of the responses of the individual optimizers is presented. It shows convergence to the site-wide optimum with significantly less iterations in comparison to the standard subgradient-based method for a set of case studies, including a petrochemical complex.

Keywords: Optimal resource allocation, market-based coordination, distributed optimization.

1 Introduction

In the process industries, energy and resource efficiency are a central concern, both for economic reasons and in order to reduce the carbon footprint and to ensure long-term sustainable production. Energy and resource efficiency are mainly determined by the design of the plants, but the way
in which the plants are operated has a significant influence as well. Model-based real-time optimization (RTO) [see 15, and references therein] or model-based optimizing control [8] are increasingly applied to improve plant operations in the presence of disturbances, changing feeds and changing demands. However, the scope of these technologies is still restricted to single plants within a larger site. In many cases, as in large chemical sites, the plants are connected by networks of resources, e.g., steam on different pressure levels, hydrogen, or intermediates. Then the individual optima may not be the best operating points for the overall site as these do not necessarily respect the site-wide constraints on the availability of the shared resources. Hence, a site-wide optimization is desired to solve the arising resource allocation problem, but often cannot be realized mainly due to following reasons:

- the overall problem becomes too large to be solved efficiently,
- centralized solutions are vulnerable to missing information and to failures of the algorithm or computer system,
- the plants may belong to different business units or even to different companies, which do not want to share the plant models and constraints and the economic performance indicators.

Therefore, distributed and hierarchical optimization methods are of strong interest (see [2]). The coordination of independently operated units that contribute to the overall performance of a technical or socio-technical system is a generic problem in the area of so-called cyber-physical systems of systems (see, e.g., [9]).

Applications of distributed and hierarchical optimization can be found in various fields such as demand-side management of electrical grids [4], smart homes and distributed charging of electric vehicles [11], distributed estimation [24], or distributed coordination of multi-agent systems of unmanned ground, aerial, and underwater vehicles [3]. Different techniques can be found in the literature for distributed and hierarchical optimization [14] depending on the degree of autonomy of the constituent systems, e.g., to which extent the systems assign some of their degrees of freedom to a local management (optimization), and based on the overall system architecture, e.g., whether there exists a central entity that coordinates the individual decisions by issuing some global reference signals. Here two main classes are represented by the algorithms that build upon primal decomposition, which are sometimes referred to as auction-based algorithms, and the techniques that use dual (Lagrangian) decomposition, which are usually referred to as price-based methods. A common feature among these techniques is that they employ a higher-level coordinator that steers the decisions of the subsystems on the lower level by issuing the commands, i.e., assigning the values of some coordination variables, which are then incorporated into individual optimization problems of the subsystems.

Using the principle of the primal decomposition [17], the coordinator allocates the amounts of resources to individual subsystems such that the overall balance of the shared resources is fulfilled. In order to arrive at the optimal resource allocation, the coordinator collects the information about the price which the subsystems are willing to pay for the allocated amounts of resources and uses this information to generate a new allocation. For the practical realization of this iterative process, the coordinator needs relatively detailed knowledge about the individual constraints of the constituent systems such that the generated resource allocation is feasible for the individual units. For instance, a unit cannot be commanded to generate a higher amount of an intermediate than its production capacity allows.

The technique of price-based coordination has been identified as a promising candidate for the realization of distributed management in process systems [12, 21]. As opposed to primal decomposition the confidentiality of the information is ensured among the constituent plants in the practical realizations as the feasibility of the individual decisions is not of a concern. This is due to the coordination principle, where the coordinator issues prices of the shared resources that enter the economic objectives of the individual subsystems. This way the coordinator stimulates (or discourages) the production (or consumption) of the shared resources by individual units. By iterative price setting, the coordinator then steers the individually optimal decisions of the subsystems towards the balance of the shared resources.

An obstacle for the industrial application of the price-based coordination technique is the slow convergence rate of the scheme that is frequently observed [16]. The objective function can be augmented by a quadratic term for the violation of the coupling constraints which convexifies the cost function and improves the speed of convergence. Then the Lagrangian can no longer be trivially decomposed and additional coordinating variables have to be introduced and the Alternating Direction Method of Multipliers (ADMM, Boyd et al. [2]) can be applied [20]. However, convergence of the ADMM method for the example considered in Stojanovski et al. [20] and below is also slow.

A comparison of available state-of-the-art distributed algorithms is given in [13]. A significant improvement in the rate of the convergence of the standard price-based...
coordination can be achieved when the coordinator possesses knowledge on the sensitivities of the subsystems to the price of the shared resources \([5, 16, 18]\). The coordinator’s knowledge is, however, limited if the subsystems do not want to share information about their optimization models and constraints. Thus, in order to improve the convergence, the coordinator needs to build the required knowledge from the observed data, i.e. the responses of the subsystems to the price signals. In order to address this issue, in this contribution we present a new scheme for price-based coordination in large-scale systems which are coupled by flows of resources and where optimizers of the individual subsystems are in operation and react to the price signals. The method builds upon techniques for estimating surrogate models from the domain of derivative-free optimization.

The rest of the paper is structured as follows. First, price-based coordination is briefly explained and linked to the general formulation of an optimal resource allocation problem. Then a new price update scheme is introduced and case studies of different complexity are formulated. In the simulation studies the performance of the new scheme is compared to the classical subgradient-based approach. Finally, conclusions are drawn and challenges are identified.

2 Price-based coordination

Price-based schemes imitate the process of a Walrasian auction \([22]\). In a Walrasian auction a central coordinator, called the auctioneer, and different competing agents take part. The agents can be producers or consumers of goods that are traded at the auction. The goal of the auctioneer is to find the equilibrium price that clears the micro market, i.e., the market comprising all present agents. If the equilibrium price is assigned to the traded goods, the supply of goods matches exactly the demand and no agent can further increase its profit.

The procedure to find the equilibrium price iteratively by dynamic pricing based on the demand and supply responses of the agents only is referred to as the tâtonnement process. The central coordinator evaluates the demand and supply of all agents for a certain price. If the supply of a certain good exceeds the demand, the coordinator will decrease the price for this good, providing an incentive to demand more and to supply less. Vice versa, if the demand is higher than the supply, the price is increased to encourage the agents to increase the supply or to demand less.

Figure 1: Typical management structure of a large industrial complex with distributed optimizers (Opt.\(i\)) and (partially) autonomous plants (Units\(i\)), coupled by shared resources. The arrows indicate which different pieces of information are exchanged in price-based coordination algorithms \(R\): shared resources, \(\lambda\): prices for shared resources, \(u\): manipulated variables, \(J\): objective function values. Price-based algorithms reflect the existent management structures and mimic the decision making process in the iteration procedure.

2.1 Price-based coordination for optimal resource allocation

The setup of a central coordinator with different competing agents can be applied to industrial production complexes, where different production plants are combined to highly integrated production sites. It is common that the production site has a central management, while the subsystems (plants or business units) have (partial) managerial autonomy with respect to their individual operation (see Figure 1). The different individual production plants (Units\(i\)) that are optimized by the optimizers (Opt.\(i\)) are strongly connected by networks of shared resources, which link the plants on the site. For instance, a power plant located in a chemical production complex feeds steam at different pressure levels to the steam networks of the site. From these networks the processing plants can take steam, e.g., to heat up the content of a reactor. The individual plants can also feed steam into the networks, which can, e.g., be produced by exothermic reactions.

Since for most of the shared resources there is only limited or no storage capacity at all, it is necessary to balance the quantities that are fed to a network and the quantities that are taken from it. The site should then be optimized subject to the constraint of balanced networks. The objectives of site-wide optimization can be manifold and are typically expressed by economic cost functions or other performance indicators, e.g., CO\(_2\) footprint. We as-
sume that the following general optimization problem has to be solved
\begin{align}
\min_{u_i \in U_i} & \sum_{i=1}^{n} J_i(u_i), \\
\text{s.t.} & \sum_{i=1}^{n} R_i(u_i) + R_e = 0,
\end{align}

where $J_i : \mathbb{R}^{n_{i}} \to \mathbb{R}$ denotes the objective function of the plant $i$ that is a function of the vector of individually manipulated variables $u_i \in \mathbb{R}^{n_{i}}$ whose values are constrained by the set $U_i$. The site-wide objective is the sum of the individual objective functions $J_i$ and can be decoupled and optimized in a distributed fashion. However, the decisions of the plants are coupled by the constraint (1b), which sums up the $n$ shared resource utilization vectors $R_i : \mathbb{R}^{n_{i}} \to \mathbb{R}^{n_e}$ (see Section 3.1) and the external flows $R_e$ that leave or enter the balance space (see Figure 1). Such an external flow can be for example electricity that is sold to the grid or an incoming utility stream from adjacent production complexes. The coupling constraint is satisfied if the sum of all the vectors vanishes. A supply is denoted with a negative sign ($R_{ij} < 0$), whereas a demand is a positive quantity ($R_{ij} > 0$), where $j$ is the index of a particular shared resource (e.g., $R_{3,3} = -2.5 \text{ t h}^{-1}$ denotes a supply of 2.5 tons per hour of resource 4 by plant 3).

The central site management has only limited influence and knowledge about the individual units and optimizers, but it has the ability to adjust the transfer prices $\lambda$ for the shared resources in the distribution networks. While keeping the autonomy of the individual production plants, the central site management can thus give incentives to change the individual operation in order to contribute to balanced networks in the production site.

In order to realize the price-based coordination the objective functions on the lower (individual) level of the plants (1a) are augmented by the product of the Lagrange multipliers $\lambda \in \mathbb{R}^{n_e}$ and the residual of the coupling constraint to form the Lagrange function
\begin{equation}
\mathcal{L}(u, \lambda) = \sum_{i=1}^{n} J_i(u_i) + \lambda^T \left( \sum_{i=1}^{n} R_i(u_i) + R_e \right),
\end{equation}
that needs to be minimized [1]. The Lagrange multipliers here represent the penalty parameters for the unit increase (or decrease) of the consumption (or production) of the resources. Thus, they can be interpreted as transfer prices by which the individual plants are either penalized for taking shared resources from the network, or rewarded for the production of shared resources [2].

Since the coupling constraint is relaxed, the problem becomes separable. Consequently, each subsystem can optimize its own cost function taking into consideration the current transfer price $\lambda$ that is set by the coordinator
\begin{equation}
\min_{u_i \in U_i} J_i(u_i) + \lambda^T R_i(u_i).
\end{equation}

Here we assume that
- the functions $J_i(u_i)$ are strictly convex,
- the functions $R_i(u_i)$ are affine,
- the individual constraints are satisfied with strict inequalities at the site-wide optimum,
- and that only the coupling constraint determines (constrains) the optimum.

These assumptions are often fulfilled in industrial practice where the availability of the shared resources is the limiting factor of the production. Under these assumptions there exists a unique optimizing pair $(u^*, \lambda^*)$ that satisfies [1]
\begin{align}
0 &= \nabla_u \mathcal{L}(u^*, \lambda^*), \\
0 &= \nabla_\lambda \mathcal{L}(u^*, \lambda^*), \\
0 &< \nabla_u^2 \mathcal{L}(u^*, \lambda^*).
\end{align}

One of the consequences of these optimality conditions is that the Lagrangian is locally quadratic around the optimum.

### 2.2 Price update step

While on the lower level of the management structure individual optimization problems are solved for a current price vector $\lambda^k$, on the upper (coordinator) level the price vector is updated in an iterative procedure, with $k$ as the iteration index. The dual function of (1) is, under the aforementioned assumptions, a strictly concave function that is defined as
\begin{equation}
d(\lambda) := \min_{u \in U} \mathcal{L}(u, \lambda).
\end{equation}

This function has to be maximized in order to find a saddle-point of Lagrangian given by conditions (4a) and (4b). Because of this, the next price vector $\lambda^{k+1}$ can be found by moving along the direction of the gradient of the dual function
\begin{equation}
\lambda^{k+1} = \lambda^k + \alpha^k \nabla d(\lambda^k),
\end{equation}
where $\alpha^k \in \mathbb{R}$ is the step size parameter. If local constraints are active, the dual function may not be differentiable. To cover such cases the update step is often defined using the subgradient method that is a generalization of the gradient-based search (see [19] for details). The
subgradient-based update is realized as follows

\[ \lambda^{k+1} = \lambda^k + \alpha^k \mathcal{G}(\lambda^k), \]  
where a subgradient \( \mathcal{G} \) is given by the evaluation of the network balance incorporating the optimal responses of the subsystems. A subgradient for the resource allocation problem can be inferred from the dual function (5) as

\[ \mathcal{G}(\lambda^k) = \sum_{i=1}^{n} R^k_i (u^*_i) + R_x, \]  
where \( u^*_i \) denotes the optimal operating point of subsystem \( i \) for the price vector \( \lambda^k \). The price vectors are iteratively updated until the network balance is achieved. At this point a subgradient vanishes and the optimal price vector \( \lambda^* \) has been found.

The price-based coordination is guaranteed to converge under the aforementioned assumptions for sufficiently small \( \alpha \) [1]. On the other hand, \( \alpha \) should be sufficiently big to achieve as fast convergence as possible, hence defining a trade-off. An extensive discussion about different choices of \( \alpha \) and convergence results can be found in Bertsekas [1]. Commonly used proposals are a fixed step size parameter, a fixed step length or diminishing step sizes, chosen off-line, before the start of the coordination.

Since the appropriate choice, both in terms of convergence rate and robustness, is not trivial, investigating alternative ways to find the optimal price \( \lambda^* \) is an interesting field of research. A novel approach is discussed in Section 4.

3 Problem formulation

In this chapter a generic formulation for a site-wide steady state resource allocation problem with stationary linear plant models is presented. Such models are commonly used in industry on a planning level and for the long-term coordination of shared resources. We note that the novel price update methodology can easily be generalized to linear dynamic models.

3.1 Mathematical modeling of the plants

The mathematical models of the plants are assumed to be affine mappings. The state vector \( x_i \in \mathbb{R}^{n_i} \) of plant \( i \) can be computed as a function of the manipulated variables \( u_i \in \mathbb{R}^{n_u} \) by

\[ x_i = M_{ux,i} u_i + V_{x,i}, \]  
where \( M_{ux,i} \in \mathbb{R}^{n_x \times n_u} \) is the linking matrix and \( V_{x,i} \in \mathbb{R}^{n_x} \) defines a constant offset of the model. The product vector \( y_j \in \mathbb{R}^{n_j} \) of plant \( i \) is related to the state vector by

\[ y_j = M_{xy,j} x_i, \]  
with the matrix \( M_{xy,j} \in \mathbb{R}^{n_y \times n_x} \). The shared resources contain contributions from the inputs and states. Hence, the resource utilization vector \( R_i \in \mathbb{R}^{n_R} \) that contains the production as well as consumption, is determined by the following equation

\[ R_i = M_{xR,i} x_i + M_{uR,i} u_i, \]  
with \( M_{uR,i} \in \mathbb{R}^{n_u \times n_u} \) and \( M_{xR,i} \in \mathbb{R}^{n_x \times n_x} \). As mentioned earlier, negative entries of a vector \( R_i \) denote the production of a shared resource, while positive values indicate consumption. If a shared resource is neither consumed nor produced, or the plant is not connected to a specific network, the entry is set to zero.

An example of such a plant model is a plant \( P \) that processes raw material \( m_{30} \) using 30 bar steam \( m_{30} \) to produce the product \( m_{5} \). It sends 5 bar steam \( m_{5} \) to the shared resource networks and receives 30 bar steam. The aforementioned formulation with \( u_p = (m_{30}, m_{5})^T \) in this case is

\[ x_p = \begin{pmatrix} m_5 \\ m_{50} \end{pmatrix} = \begin{pmatrix} 7.8 & 0.0 \\ 0.0 & 2.6 \end{pmatrix} u_p + \begin{pmatrix} 3.4 \\ 0.0 \end{pmatrix}, \]  
\[ y_p = m_{50} = \begin{pmatrix} 0 & 1 \end{pmatrix} x_p, \]  
\[ R_p = \begin{pmatrix} -m_5 \\ m_{30} \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} x_p + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} u_p. \]  
The parameters used in (12) are only exemplary.

3.2 Optimal resource allocation problem

In the following it is assumed that the cooperative resource allocation problem can be formulated as a constrained separable quadratic program. This assumption is stronger than the previously used one of strictly convex objective functions of the individual subsystems. On the other hand, the formulation using quadratic objective functions elucidates the following discussions and it can be regarded as a quadratic approximation of the problem (1) that becomes exact around the optimum. Then the linear terms in the objective function represent the plant economics, which are linear by nature, e.g., prices multiplied by mass flows, and the quadratic terms result from tracking certain production targets, which can be part of contracts. The math-
ematically program – neglecting constant terms in the objective function – can then be formulated as

$$\min_{u_i \in U_i, y_i \in Y_i} \sum_{i=1}^{n} \frac{1}{2} u_i^T P_i y_i + \frac{1}{2} u_i^T \hat{Q}_i u_i + p_i^T y_i + \hat{q}_i^T u_i$$

(13a)

subject to

$$M_{xy,i} x_i = y_i,$$  \hspace{1cm} (13b)

$$M_{ux,i} u_i + V_{x,i} = x_i,$$  \hspace{1cm} (13c)

$$M_{xr,i} x_i + M_{ur,i} u_i = R_i(x_i, u_i),$$  \hspace{1cm} (13d)

$$\sum_{i=1}^{n} R_i(x_i, u_i) + R_e = 0,$$  \hspace{1cm} (13e)

where the matrices $P_i \in \mathbb{R}^{n_i \times n_i}$ and $\hat{Q}_i \in \mathbb{R}^{n_i \times n_i}$ are symmetric and positive definite ($P_i, \hat{Q}_i > 0$). The relations (13b)–(13d) contain the stationary model equations of the plants, where $x_i \in \mathbb{R}^{n_i}$ are the states of plant $i$ and $y_i \in \mathbb{R}^{n_i}$ are the products. The constraint (13e) is the network balance, where $R_e \in \mathbb{R}^{n_e}$ is the offset that incorporates fixed resource flows over the boundaries of the production complex.

Since the states $x_i$ and the products $y_i$ are affine functions of the manipulated variables, the problem can be rewritten in the following form

$$\min_{u_i \in U_i, y_i \in Y_i} \sum_{i=1}^{n} \frac{1}{2} u_i^T Q_i u_i + q_i^T u_i,$$  \hspace{1cm} (14a)

subject to

$$\sum_{i=1}^{n} R_i(u_i) + R_e = 0,$$  \hspace{1cm} (14b)

where the $Q_i$ and $q_i$ are defined by the following expressions

$$Q_i = \hat{Q}_i + M_{ux,i}^T M_{xy,i} P_i M_{xy,i} M_{ux,i},$$  \hspace{1cm} (15a)

$$q_i = \hat{q}_i + (V_{x,i} M_{xy,i} P_i + p_i) M_{xy,i} M_{ux,i}.$$

(15b)

The constant terms are dropped since they do not influence the result of the optimization.

### 4 Novel price update strategy

As previously mentioned it is of interest to formulate a price update strategy that finds the equilibrium price in few iterations. With the assumption of quadratic cost functions, affine models, and the individual constraints being inactive at the optimum, a relationship can be established between the residual of the network balance (coupling constraint) and the price vector $\lambda$. In Wenzel et al. [23] it is shown that the residual of the network balance can be expressed as a quadratic function of the price vector. The relation can be written as

$$f_\lambda(\lambda) = \left\| \sum_{i=1}^{n} R_i(u_i^*) + R_e \right\|_2^2 = \lambda^T H\lambda + s^T \lambda + v.$$  \hspace{1cm} (16)

This relation is locally valid at the optimum also in the more general case of strictly convex cost functions, since the first- and second-order conditions of optimality (4c) state that the objective function can be represented as a quadratic function near the optimum. The new update strategy exploits this relation to find the optimal price by minimizing $f_\lambda : \mathbb{R}^{n_e} \rightarrow \mathbb{R}$, which can be done in one step if $H$ and $s$ are known. In practice, $H \in \mathbb{R}^{n_e \times n_e}, s \in \mathbb{R}^{n_e}$ and $v \in \mathbb{R}$ are not available, since they incorporate information about the individual cost functions. Thus, the function $f_\lambda$ has to be identified from the information that is available to the coordinator.

The idea is to probe the subsystems with a sufficient number of price vectors and to record the responses, i.e., the corresponding residuals of the network balance. For a quadratic relationship (16) a sufficient amount of measurements is

$$n_{QA,\text{min}} = \frac{(n_R + 1)(n_R + 2)}{2},$$  \hspace{1cm} (17)

with $n_R$ being the number of shared resources. If enough points are available, $H, s$ and $v$ can be determined. The procedure is inspired by derivative-free optimization techniques (DFO) that build surrogate models based on sampled function values [7]. If the points that are used to approximate the quadratic function are well distributed and not too close together, i.e., they form a well-poised set, then the fit will be good even in the presence of small errors in the values. If the points are linearly dependent or they are too close together, the quadratic approximation can be imperfect, or it can even fail. Hence, the strategy for the selection of probing points for the quadratic approximation has to be chosen carefully. In Gao et al. [10] a screening algorithm is proposed that is able to select a well-poised set of points for a robust and reliable quadratic approximation. The proposed algorithm was developed for the estimation of reliable plant gradients in the Modifier Adaptation method, but it can also be used in the framework of price-based coordination for selecting a suitable set of points for the quadratic approximation of (16).

If the assumptions formulated for (14) and the assumption of inactive individual constraints do not hold, the relation of the network residual and the prices may not be quadratic. In this case the optimum cannot be found within one iteration. However, it was shown in [6] that a quadratic approximation model can approximate...
Figure 2: Block diagram of the novel coordination procedure. SG: subgradient-based price update (7), RQA: recursive quadratic approximation price update (18).

any convex differentiable function with arbitrary precision. If either the chosen set of points is not well-poised and the fit is imperfect or the optimal price is not found because of a non-quadratic relation of the network residual, the quadratic fitting is repeated recursively (recursive quadratic approximation, RQA). Thus, to find the optimal price, the strategy shown in Figure 2 is applied to solve the optimal resource allocation problem. The algorithm is initialized with the necessary settings as, e.g., termination tolerances, maximum number of iterations, and number of required probing points $n_{QA}$. First, an initial price vector $\lambda^0$ is sent to the individual subsystems with their attached individual (local) optimizers ($LO_i, \forall i$). The individual subsystems respond with their shared resource utilization $R^k_i$, based on the current price vector. The responses of all subsystems are collected and the residual is evaluated. If the convergence criterion is not met, the next price vector $\lambda^{k+1}$ is determined. As long as less probing points than $n_{QA}$ are available, the price update is done using the classical subgradient-based update (7). Here, $n_{QA}$ is, at least equal to the number of points that is needed to fit a quadratic function, $n_{QA_{min}}$, but can also be chosen larger to collect more initial data before the first quadratic approximation. The initial points can also be collected by alternative strategies. However, using subgradient-based price updates is a reasonable choice if no information about the local subsystems is available in the initial phase. As soon as $n_{QA}$ points have been collected, the first quadratic approximation $f^k_r^{n_{QA}}$ can be determined. The next price is then found by optimizing $f^k_r^{n_{QA}}$

$$\lambda^{k+1} = \arg\min_{\lambda \in \Lambda^k} \lambda^k_r(f_r^k(\lambda)), \quad (18)$$

where $k$ is the iteration index and $\Lambda^k$ is a limited search space for the next price vector (a box constraint in the simplest case).

If more than $n_{QA}$ points are available, the sets of points are selected from the available data by the screening algorithm described in Gao et al. [10]. The next quadratic approximation is then based on the selected sets.

5 Case studies

In this section we present case studies of different complexity. First, resource allocation problems are defined that consist of unconstrained individual plants that are coupled by different numbers of utility networks. Secondly, a case study is presented that is based on the structure of the petrochemical site of INEOS in Köln as described by Stojanovski et al. [20].

5.1 Coordination of unconstrained individual production plants

First, we consider three unconstrained individual production plants with individual optimizers as shown in Figure 3. The individual cost functions are quadratic functions that incorporate linear economic and quadratic tracking terms, which result from casting the production demand into a reference tracking problem. The individual cost functions have the following form

$$J_i(u_i) = p_{1,i} \cdot u_{i,1} + p_{2,i} \cdot u_{i,2} + \frac{(u_{i,1,\text{ref}} - u_{i,1})^2 + (u_{i,2,\text{ref}} - u_{i,2})^2}{\text{sales and raw material costs}} + \frac{(u_{i,1,\text{ref}} - u_{i,1})^2 + (u_{i,2,\text{ref}} - u_{i,2})^2}{\text{input tracking}}, \quad i \in \{A, B, C\}, \quad (19)$$

where $u_i \in \mathbb{R}^2$ contains the manipulated variables, $u_{i,\text{ref}}$ are references and $p_{ij}$ are the prices that are assigned to bought or sold quantities, i.e., produced and consumed shared resources.

Following the formulation of the optimization problem (13) and the expression for the shared resources (11), the optimal resource allocation problem for three
Figure 3: Case study consisting of three plants with unconstrained individual quadratic cost functions. Two resources ($R_1$ and $R_2$) are shared among the plants. Additionally, two export streams $R_{1,e}$ and $R_{2,e}$ leave the networks, which have fixed values. In the contour plots the dots indicate the uncoordinated (decentralized) optima of the plants (i.e., optimal with respect to the individual cost functions, but not respecting the coupling constraint), while the diamonds show the optimal individual resource utilizations that fulfill the coupling constraint.

quadratic subsystems and two shared resources can be written as

$$
\min_{u_i,\lambda} \sum_{i=1}^{n} \frac{1}{2} u_i^T Q_i u_i + q_i^T u_i ,
$$

(20a)

$$\text{s.t.} \sum_{i=1}^{n} R_i(u_i) + R_e = 0 ,
$$

(20b)

$$R_i(u_i) = M_{uR,i} u_i ,
$$

(20c)

where $n = 3$ and $R_e \in \mathbb{R}^2$ denotes the vector of exported quantities. The model parameters are given in Table 1. The central formulation of the optimization problem can be distributed and the following individual Lagrangian functions result

$$L_i(u_i, \lambda) = \frac{1}{2} u_i^T Q_i u_i + (q_i^T + \lambda^T M_{uR,i}) u_i ,
$$

(21)

which give the optimal resource utilization of subsystem $i$ as a function of the transfer price vector $\lambda$

$$R_i(\lambda) = M_{uR,i} \arg \min_{u_i} L_i(u_i, \lambda)
$$

$$= -M_{uR,i} Q_i^{-1} (q_i + M_{uR,i}^T \lambda) .
$$

(22)

As extensions to the coordination of three production plants, the problem size is increased to 10, 100, and 1000 subsystems linked by 2, 3, and 5 shared resources to study the scalability of the proposed price update scheme. The formulation of the problems is as (20). The symmetric positive definite matrices $Q_i \in \mathbb{R}^{n_i \times n_i}$ with $n_i \in \{2, 3, 5\}$ are randomly generated using samples from a normal distribution with zero mean and a standard deviation $\sigma$ of 1 ($N(\mu = 0, \sigma^2 = 1)$). In a similar way the vectors $q_i \in \mathbb{R}^{n_i}$ are sampled from a normal distribution with zero mean and a standard deviation $\sigma$ of 10 ($N(0, \sigma^2 = 100)$). The matrices $M_{uR,i} \in \mathbb{R}^{n_i \times n_R}$ have zero off-diagonal elements and the main diagonal elements were sampled from a uniform distribution $U(\mu, 1)$.

Table 1: Model parameters for the coordination of three unconstrained quadratic subsystems. Negative prices denote sales prices.

<table>
<thead>
<tr>
<th>Plant</th>
<th>$p_{1,i}$</th>
<th>$p_{2,i}$</th>
<th>$u_{1,i,\text{ref.}}$</th>
<th>$u_{2,i,\text{ref.}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>-10</td>
<td>8</td>
<td>30</td>
<td>18</td>
</tr>
<tr>
<td>B</td>
<td>12</td>
<td>-14</td>
<td>20</td>
<td>12</td>
</tr>
<tr>
<td>C</td>
<td>11</td>
<td>-13</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

$$R_e = \begin{pmatrix} -5 \\ 6 \end{pmatrix}^T$$

$$M_{uR,A} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$M_{uR,B} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$M_{uR,C} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$Q_i = 2 \cdot I$$

$$q_i = \begin{pmatrix} p_{1,i} - 2 u_{1,i,\text{ref.}} \\ p_{2,i} - 2 u_{2,i,\text{ref.}} \end{pmatrix}$$

(23)

5.2 Steam management in a petrochemical production site

This case study is based on a simplified model of the integrated petrochemical production site of INEOS in Köln as described by Stojanovski et al. [20].
5.2.1 Overview of the case study

The resource allocation problem comprises a power plant, a cracker, and an ammonia plant that share steam at different pressure levels (5 bar and 30 bar) via two shared resource networks as illustrated in Figure 4. The structure of the case study is comparable to the coordination task of the three unconstrained production plants in Section 5.1. However, the individual plants in the industrial case study are constrained by bounds (23a) on input (manipulated) variables, states and outputs. These constraints represent operational constraints of the equipment, safety considerations and contracts with customers or producers on the site. Additionally, the set of input and output variables is different from the set of shared resource variables. In the following a brief description of the single plants is given.

The main purpose of the power plant is to burn waste gases of adjacent production processes on site or fuel gas in order to produce electricity and steam on two different pressure levels which are needed by the industrial complex. The manipulated variables considered here are the fuel gas consumption \( \dot{m}_{FG} \), and the power plant performance factor \( PP_{perf} \), which determines the ratio of the two steam levels produced by this plant. The electrical power is considered as a product that can be sold, while the two steam flows are sent to the shared resource networks for 5 bar and 30 bar steam.

The cracker represents the heart of the petrochemical production site. The incoming naphtha stream \( \dot{m}_{LD} \) is cracked in order to produce different petrochemical intermediates, ethylene and propylene with flowrates \( \dot{m}_{C2} \) and \( \dot{m}_{C3} \) being the most important ones. In addition to the input of naphtha, the severity of the cracker \( \alpha \) (i.e., the ratio of C2/C3) and the input stream of an additional 105 bar steam stream can be manipulated.

In the ammonia plant hydrogen and fuel gas are used to produce ammonia, the main product of this plant. The manipulated variables are the natural gas intake \( \dot{m}_{NG} \), the hydrogen intake \( \dot{m}_{H2} \), and the deionized water consumption \( \dot{m}_{DW} \). The plant uses steam from the 5 bar network and produces 30 bar steam.

5.2.2 Resource allocation problem

If the network balance is neglected, the plants (power plant, cracker, ammonia plant) solve the following individual optimization problems in parallel in a distributed fashion

\[
\begin{align*}
\min_{u_i \in U_i, x_i \in X_i, y_i \in Y_i} \frac{1}{2} y_i^T P_i y_i + \frac{1}{2} u_i^T Q_i u_i + p_i^T y_i + q_i^T u_i \\
\text{s.t. } M_{xy,ij} x_i = y_i, \\
M_{ux,i} u_i + V_{xj} = x_i, \\
M_{yR,ij} x_i + M_{uR,i} u_i = R_j(x_i, u_i),
\end{align*}
\]

where (23b) and (23c) are the model equations. For the sake of clarity, (23b) and (23c) are employed to formally eliminate \( x_i \) and \( y_i \) to yield the objective functions \( f_i(u_i) \), and the constraint functions \( R_j(u_i) \), \( \forall i \). The optimization problem that has to be solved for the overall production
complex includes the coupling constraint of balanced networks.

\[
\min_{u_i \in U_i} \sum_{i=1}^{n} f_i(u_i) \\
\text{s.t. } \sum_{i=1}^{n} R_i(u_i) + R_e = 0.
\]

(24a) (24b)

By Lagrangian relaxation, the coupling constraint is introduced into the objective function and a separable Lagrangian function is defined

\[
\mathcal{L}(u, \lambda) = \sum_{i=1}^{n} f_i(u_i) + \lambda^T \sum_{i=1}^{n} R_i(u_i),
\]

where the exported or imported quantities \(R_e\) can be neglected, since they are constant. The optimal shared resource utilization of an individual plant for a given price vector \(\lambda\) can then be formulated as

\[
R_i(\lambda) = M_{uR_i} \arg\min_{u_i \in U_i} \mathcal{L}(u_i, \lambda).
\]

(25) (26)

6 Simulation results

In the following section simulation results of the two described case studies are provided. For both case studies price-based coordination is applied, comparing the classical subgradient-based price update (SG) with the novel approach of recursive quadratic approximation (RQA) as described in Section 4.

In addition to the evolution of the price-vectors \(\lambda^k\) and the squared 2-norms of the network residual versus the number of iterations, an analysis of the response surfaces of the network residual as a function of the price vector is performed. The resulting heat and contour plots provide insights into the properties of the problems and the behavior of the solution algorithm.

6.1 Results for the coordination of unconstrained individual production plants

For the coordination of three unconstrained subsystems as defined in Section 5.1 the simulation results are shown in Figure 5a and 5b. The dashed lines show the evolution of the subgradient-based update for a fixed parameter \(\alpha^k = 0.25\) – it was chosen to provide a fast, but non-oscillatory convergence – and a termination tolerance of \(\epsilon = 10^{-4}\). It can be seen that after approximately ten iterations the prices are already near the optimal price. However, it takes another ten iterations for the algorithm to converge, since the residual is decreasing linearly. So the subgradient-based algorithm provides a slow movement towards the optimal price in its vicinity. The latter is a well-known feature of the first-order search.

The RQA builds the first quadratic approximation (model) after six iterations, resulting first in updated prices that worsen the network balance. Nevertheless, it can be observed that after these few iterations the RQA converges exactly to the optimal price vector. The movement in the \(\lambda\)-space can be interpreted as exploration phase in which the algorithm collects necessary information on the network mappings to set up a well-poised set with linearly independent points for the approximation (see Section 4).

The results for the extensions of the problem are shown in Table 2. For combinations of 10, 100, and 1000 subsystems with 2, 3 or 5 shared resources the number of iterations needed to converge is indicated for the subgradient-based updates and the quadratic approximation method. Additionally, for every scenario the chosen value of \(\alpha\) is given, which was adjusted to have a minimum number of iterations for the subgradient-based updates (i.e., slightly higher values of \(\alpha\) lead to diverging behavior of the subgradient-based updates). It can be seen that for all scenarios the new method outperforms the subgradient-based updates in terms of required iterations, even for a well-tuned value of \(\alpha\). A remarkable result is how the required iterations scale with the number of shared resources. While, e.g., the balancing of 1000 subsystems with 2 resources requires 19 iterations with the subgradient-based method, but already 161 to balance five resources, the number of iterations with the proposed method only increases from 13 to 59. It can be seen for the quadratic approximation that the number of iterations stays approximately constant for the same number of shared resources regardless of the number of plants.

The distributed optimization converged to the same solutions as a centralized method. The centralized method, however, needed significantly shorter computation time to converge. This fact points out the advantages of structure-exploiting centralized optimization if it is realizable.

6.2 Results for the steam management in a petrochemical production site

The industrial case study of three individually constrained production plants shows a higher sensitivity to transfer prices in comparison to the unconstrained plants. Consequently, the price update parameter \(\alpha^k\) is chosen to be
Figure 5: Simulation results for the coordination of three unconstrained individual production plants.

Table 2: Number of iterations for the optimization of 10, 100, and 1000 subsystems with $n_R \in \{2, 3, 5\}$ shared resources needed to converge to the threshold of $\varepsilon_r = 10^{-4}$ for the subgradient-based updates (SG) and the recursive quadratic approximation (RQA). The values for $\alpha$ in (7) were tuned for each problem to achieve the minimum number of iterations for the subgradient-based updates.

<table>
<thead>
<tr>
<th>$n_R$</th>
<th>10 subsystems</th>
<th>100 subsystems</th>
<th>1000 subsystems</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SG</td>
<td>RQA</td>
<td>$\alpha$</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>9</td>
<td>0.25</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>18</td>
<td>0.21</td>
</tr>
<tr>
<td>5</td>
<td>105</td>
<td>56</td>
<td>0.044</td>
</tr>
</tbody>
</table>

2.5 \cdot 10^{-3}$ – higher values led to diverging behavior – and the termination tolerance of the squared 2-norm of the residual of the network $\varepsilon_r$ is set to $10^{-2}$, which is a reasonable value for the real plant. The allowed price change constraint for the RQA (see (18)) is chosen to be a fixed-size trust-region $\Lambda^k = \|\lambda^{k+1} - \lambda^k\|_2 \leq 10.0$. The simulation results are shown in Figure 6a and 6b. It can be seen that the subgradient-based updates improve the network balance from the initial point on, but there are jumps in the prices around iteration 60 and 145. This behavior is caused by the individual constraints of the subsystems that are not visible to the coordinator and which lead to a high sensitivity of the response of the network balance (see Section 6.3). After the jumps the subgradient-based prices recover and slowly improve the network balance towards the termination tolerance. But within the $k_{\text{max}} = 200$ iterations the subgradient-based update is not able to balance the networks to the desired termination tolerance $\varepsilon_r$.

In contrast, the RQA shows a much better rate of convergence in this simulation. Initially, the movement in the
price-space seems to be undirected, but the collected information leads to the determination of the equilibrium prices after less than 40 iterations. A possible explanation for the large number of iterations that is still needed can be found in the problem structure. While in the unconstrained case in the first case study the shared resources can be adjusted independently from each other, in the case study of INEOS in Köln there exist an interdependency of the production/consumption of the shared resources that is hidden to the coordinator.

6.3 Analysis of the response surfaces

A deeper insight into the problem structure can be gained by plotting the response surfaces, i.e., the network residual as function of the prices, for both case studies as shown in Figures 5c and Figure 6c. On the horizontal and vertical axes the two prices for the two shared resources are plotted as coordinates. The heat plot shows the constraint violation, i.e., the network residual (16). For the case study of unconstrained individual plants a quadratic mapping can be identified as in (16). In case of the individually constrained case study the situation is different. While some regions can be described by a quadratic mapping, the influence of the individual constraints of the subsystems is clearly visible. The overall site optimum is determined by the individual constraints of the subsystems, which is frequently encountered. For the price-based coordination this causes challenges. Near the optimum, the response surface is very sensitive to price changes and large violations occur when the optimal price is perturbed by just a small value.

How sensitively the networks react to small changes of the prices can be seen in the contour plots Figures 5d and 6d, in which a subset of the data of Figures 5c and 6c is shown. For the coordination of the unconstrained production plants it can be seen, that the residual of the net-
work is increasing in both directions of $\lambda$ in the same order of magnitude, i.e., the network is sensitive in both directions. The subgradient-based method and the proposed RQA are able to converge to a point within the $\varepsilon_r = 10^{-4}$ line. In contrast, for the industrial case study the network shows different sensitivities with respect to the two prices. In Figure 6d it can be seen that a change in $\lambda_1$ causes a larger constraint violation than a change in $\lambda_2$. This can be one reason for the fact that the subgradient-based updates have difficulties to converge to a point within the desired tolerance of $\varepsilon_r = 10^{-2}$.

7 Conclusions

It was shown that, under certain assumptions on the structure of the subproblems, price-based coordination is well suited to achieve site-wide optimal distribution of the shared resources among different competing plants. The methodology of price-based coordination preserves information confidentiality to a maximum extent by adjusting the transfer prices only based on the network balance and obtaining data on the resource consumption that results from values of the transfer prices from the subsystems.

The presented novel price update by recursive quadratic approximation showed a superior convergence rate in the simulation studies and is a step forward for practical implementations, where fewer iterations are key to the success of this strategy. The iterative estimation of the second-order information improves the search direction and convergence. The number of iterations scales better than for the subgradient-based update when the number of shared resources increases.

The presence of active constraints of the subsystems poses challenges if the overall site optimum is governed by the individual constraints of the subsystems, e.g., when a plant operates at its production limits. In this case the network balance can be very sensitive to price changes and coordination has to be performed with care. Therefore, the selection of the probing points for the quadratic approximation is performed by using the screening algorithm described in Gao et al. [10].

It should also be noted that the approach presented here is not limited to the application on a certain decision layer of the automation pyramid. Problems of the form (1) are present on planning, real-time optimization or advanced control levels and the price-based coordination methodology presented here can be applied as well.

8 Summary

In this paper a motivation for and an introduction to price-based coordination was given. A general formulation of the modeling of optimal resource allocation problems is provided together with a decomposition strategy where the individual optimization problems are solved in a distributed fashion with a central coordinator on the top level.

A novel price update step was presented and its performance was evaluated in terms of convergence for several case studies. In addition, challenges that occur in the situation of individually constrained subsystems were identified and discussed.

Acknowledgement: The authors gratefully acknowledge the support of the European Commission under the grant agreement number 611281 (FP7 ICT project DYMASOS, www.dymasos.eu).

References

Bionotes

Simon Wenzel, M.Sc.
Technische Universität Dortmund, Fakultät Bio- und Chemieingenieurwesen, Lehrstuhl für Systemdynamik und Prozessführung, Emil-Figge-Straße 70, 44227 Dortmund, Germany
simon.wenzel@bci.tu-dortmund.de

Simon Wenzel, M.Sc. (born 1988) received the M.Sc. degree in Chemical Engineering – Specialisation Process Systems Engineering, at TU Dortmund University in 2014. Since then he has been a research associate with the Process Dynamics and Operations Group at TU Dortmund University, working on his PhD in the field of distributed coordination and market-like algorithms.

Dr. Radoslav Paulen
Technische Universität Dortmund, Fakultät Bio- und Chemieingenieurwesen, Lehrstuhl für Systemdynamik und Prozessführung, Emil-Figge-Straße 70, 44227 Dortmund, Germany

Dr. Radoslav Paulen (born 1984) graduated at Slovak University of Technology in Bratislava (Slovakia) in 2008. He received his PhD from the same university in the field of Process Control in 2012. Since 2012, he has been with the Process Dynamics and Operations Group at TU Dortmund University. His research interests include distributed optimization and control, model-based optimizing control, and guaranteed parameter and state estimation.

Goran Stojanovski, PhD
Ss. Cyril and Methodius University in Skopje, Faculty of Electrical Engineering and Information Technologies, Rugjer Boshkovik 18, PO Box 574, 1000 Skopje, Republic of Macedonia

Goran Stojanovski, PhD (born 1984) received the Dipl.-Ing. and his PhD degree from the Ss. Cyril and Methodius University in Skopje, Macedonia in 2007 and 2013 respectively. In 2014, he joined the Process Dynamics and Operations Group at TU Dortmund University, working in the field of distributed coordination for plant-wide optimization. In April 2015, he was appointed to his present position as an Assistant Professor of Automation and Control Systems with the Faculty of Electrical Engineering and Information Technologies at Ss. Cyril and Methodius University in Skopje.

Dr.-Ing. Stefan Krämer
INEOS Köln GmbH, Alte Str. 201, 50769 Köln, Germany

Dr.-Ing. Stefan Krämer (born 1972) studied Chemical Engineering at the University of Newcastle upon Tyne (GB) and at TU Dortmund University. He joined INEOS in 2004 as an Advanced Control Engineer and ran a team of APC and DCS engineers from 2009 to 2013. As Site Energy Manager his responsibilities now include operating the site wide energy management system and coordinating site energy optimization and energy efficiency projects. He teaches Batch Process Operation as a Guest Lecturer at TU Dortmund University. He is active in the NAMUR working group on Energy Efficiency.
Benedikt Beisheim, M.Sc.
INEOS Köln GmbH, Alte Str. 201, 50769 Köln, Germany

Benedikt Beisheim, M.Sc. (born 1987) studied Chemical Engineering at the RWTH Aachen University and received the M.Sc. degree in 2014. Since 2014 he has been with INEOS in Köln, working in the energy management department. He is responsible for energy optimization projects and is involved in the EU-funded projects DYMASOS and MORE.

Prof. Dr.-Ing. Sebastian Engell
Technische Universität Dortmund, Fakultät Bio- und Chemieingenieurwesen, Lehrstuhl für Systemdynamik und Prozessführung, Emil-Figge-Straße 70, 44227 Dortmund, Germany

Prof. Dr.-Ing. Sebastian Engell (born 1954) received the Dipl.-Ing. degree in Electrical Engineering from the Ruhr-Universität Bochum, Germany, in 1978, and the Dr.-Ing. degree from the Universität Duisburg, Germany, in 1981. In 1990, he was appointed to his present position as a Full Professor of Process Dynamics and Operations with the Department of Biochemical and Chemical Engineering at TU Dortmund University. His areas of research are robust optimizing control, in particular of chemical and biochemical processes, dynamic modeling, coordination in systems of systems, and planning and scheduling under uncertainty. He is recipient of an ERC Advanced Investigator Grant and the current President of the European Control Association.