Optimization and Engineering (2020) 21:1095–1131
https://doi.org/10.1007/s11081-020-09502-1

RESEARCH ARTICLE

Deterministic global optimization of steam cycles using the IAPWS-IF97 model

Dominik Bongartz[1] · **Jaromił Najman**[1] · **Alexander Mitsos**[1]

Received: 22 November 2019 / Revised: 23 March 2020 / Accepted: 6 April 2020 / Published online: 2 May 2020

Abstract

The IAPWS-IF97 (Wagner et al. (2000) J Eng Gas Turbines Power 122:150) is the state-of-the-art model for the thermodynamic properties of water and steam for industrial applications and is routinely used for simulations of steam power cycles and utility systems. Its use in optimization-based design, however, has been limited because of its complexity. In particular, deterministic global optimization of problems with the IAPWS-IF97 is challenging because general-purpose methods lead to rather weak convex and concave relaxations, thus resulting in slow convergence. Furthermore, the original domains of many functions from the IAPWS-IF97 are nonconvex, while common global solvers construct relaxations over rectangular domains. Outside the original domains, however, many of the functions take very large values that lead to even weaker relaxations. Therefore, we develop tighter relaxations of relevant functions from the IAPWS-IF97 on the basis of an analysis of their monotonicity and convexity properties. We modify the functions outside their original domains to enable tighter relaxations, while we keep them unchanged on their original domains where they have physical meaning. We discuss the benefit of the relaxations for three case studies on the design of bottoming cycles of combined cycle power plants using our open-source deterministic global solver MAiNGO. The derived relaxations result in drastic reductions in computational time compared with McCormick relaxations and can make design problems tractable for global optimization.

Keywords Global optimization · Process design · Rankine cycle · Utility system · Thermodynamics · Water

✉ Alexander Mitsos
amitsos@alum.mit.edu

[1] Process Systems Engineering (AVT.SVT), RWTH Aachen University, Forckenbeckstraße 51, 52074 Aachen, Germany

1 Introduction

Water is one of the most important substances in energy conversion systems. Today, around three quarters of the global power generation relies on steam power cycles (IEA 2019), either in pure steam power plants or as part of combined-cycle plants. Furthermore, water is the most common heat transfer fluid for applications ranging from domestic heating to combined heat and power supply in utility systems for industrial production sites (Podolski et al. 2008). Given the widespread use of water in energy conversion systems, even small improvements in the design and operation of such systems can have a significant impact.

Despite the long industrial experience in the design of steam-based energy conversion systems, the optimal design of such systems (e.g., with respect to energy efficiency or cost) for given boundary conditions is still an active field of research. Steam cycle design is tackled with a variety of model-based approaches (Wang et al. 2019) ranging from simulation-based evaluation of designs manually derived from engineering experience through advanced exergy-based analyses for identifying promising points for improvement to optimization-based methods.

The state-of-the-art model for the required thermodynamic properties of water and steam is the IAPWS-IF97 (Wagner et al. 2000) in its revised version (IAPWS 2007a). It is recommended by the International Association for the Properties of Water and Steam (IAPWS) "for industrial use (primarily the steam power industry) for the calculation of thermodynamic properties of ordinary water in its fluid phases" (IAPWS 2007a). It was developed on the basis of the IAPWS-95 model (Wagner and Pruss 2002) with the goal of providing explicit expressions for all common calculations without iterative solution and with low computational effort.

While the IAPWS-IF97 is in fact the standard model for the simulation of steam cycles used in many simulators, its use in optimization-based approaches has been more limited since it is often considered too complex (Wang et al. 2019). Many existing optimization studies on steam power cycle or utility system design have either directly opted for a simpler model (Bruno et al. 1998; Bongartz and Mitsos 2017) or replaced the IAPWS-IF97 with a simplified surrogate model, typically polynomials of lower degree (Ahadi-Oskui et al. 2010). The latter approaches have sometimes been combined with smoothing techniques to remedy the nondifferentiabilities that occur at phase boundaries (Tică et al. 2012; Åberg et al. 2017). Those studies using the IAPWS-IF97 itself have done so employing different methods: stochastic optimizers (Nadir and Ghenaiet 2015), often coupled with a simulation software implementing the IAPWS-IF97 (Koch et al. 2007; Luo et al. 2011; Wang et al. 2014); local NLP solvers (Zebian et al. 2012); convex MINLP solvers (Savola et al. 2007; Manassaldi et al. 2011, 2016); or a combination of local and stochastic optimizers (Wang et al. 2015, 2016).

The existing optimization approaches using the IAPWS-IF97 share the limitation that the local or stochastic optimizers employed cannot guarantee to find a global optimum since the resulting optimization problems are nonconvex and can have multiple local solutions even for the simplest cycles (and even with

much simpler models) (Bongartz and Mitsos 2017). Therefore, deterministic global optimization is desirable for solving steam cycle design problems with the IAPWS-IF97. The most rigorous existing approach that we are aware of is the work of Ahadi-Oskui et al. (2006), who use a branch-and-cut approach for optimizing the design of combined cycle power plants modeled with the IAPWS-IF97. However, their approach is also partially heuristic since it relies on sampling of black box model functions (Nowak and Vigerske 2008).

In deterministic global optimization, a key challenge is the construction of tight convex and concave relaxations and range bounds of the functions used in the model (Locatelli and Schoen 2013). Although the form of the IAPWS-IF97 functions is such that they could in principle be addressed with general purpose methods for deriving relaxations such as αBB (Androulakis et al. 1995), the auxiliary variable method (AVM) (Smith and Pantelides 1997; Tawarmalani and Sahinidis 2002), or the McCormick technique (McCormick 1976), this is challenging for the following reasons: First, the resulting relaxations are expected to be rather weak given the complexity of the functions. In global optimization, weak relaxations result in slow convergence that makes larger problems intractable. Second, when using the AVM, the subproblems for computing lower bounds will get very large because many auxiliary variables will be added, again because of the complexity of the functions. To obtain tighter relaxations and reduce computational time in global optimization, Schweidtmann et al. (2019) recently proposed replacing complex thermodynamic models with artificial neural networks as surrogate models in the design of organic Rankine cycles. However, given the very high accuracy of the IAPWS-IF97 as well as its role as an established and trusted model implemented in most simulators, we would like to avoid replacing it with a surrogate model.

In this work, we retain the original model functions from the IAPWS-IF97 in those regions where they have physical meaning and construct tighter convex and concave relaxations and range bounds to speed up the global optimization of steam cycle design problems. Our approach is based on the observation that functions in thermodynamic models often exhibit useful monotonicity and convexity properties that can be exploited to derive tighter relaxations than those obtained by applying general purpose methods to their complex functional forms (Najman et al. 2019a, b).

In the following, we first provide some background on the methods employed for constructing relaxations. Next, we analyze the relevant functions from the IAPWS-IF97 model and describe the derived relaxations. Finally, we discuss the application of the relaxations to three case studies of combined cycle power plants to demonstrate the computational speedup compared with a plain application of the McCormick technique to the IAPWS-IF97. The results demonstrate that deterministic global optimization of problems with the IAPWS-IF97 can get intractable for larger problems when using McCormick relaxations as a general purpose technique, whereas larger problems can be solved with the proposed tailored relaxations. The

proposed relaxations are implemented in the MC++ library (Chachuat et al. 2015) used by our open-source global optimizer MAiNGO (Bongartz et al. 2018).[1]

2 Preliminaries and methods used

In the present work, we aim at enabling the use of IAPWS-IF97 functions as intrinsic functions in factorable programming techniques, in particular the multivariate McCormick technique. In the following, we thus first summarize the multivariate McCormick technique before discussing two methods that are used in Sect. 3.4 to derive relaxations of the IAPWS-IF97 functions, namely methods for relaxations of componentwise convex or concave functions and variants of the αBB method.

Throughout this article, we represent scalars as lowercase or uppercase letters (e.g., p or T), vectors in boldface (e.g., $\mathbf{x}$), and sets in calligraphic typeface (e.g., $\mathcal{X}$). $\mathbb{IR}^n$ denotes the set of all nonempty compact interval subsets of $\mathbb{R}^n$. The superscripts L and U denote the left and right end points of an interval, respectively (e.g., $\mathcal{X} = [x^{\mathrm{L}}, x^{\mathrm{U}}]$). Given a function $f : \mathcal{X} \to \mathcal{Y}$, we denote by $f(\mathcal{X})$ the image of $\mathcal{X}$ under f.

2.1 Multivariate McCormick relaxations

Established methods for deterministic global optimization are based on spatial branch-and-bound (B&B) (Falk and Soland 1969) and rely on the availability of valid convex and concave relaxations (Locatelli and Schoen 2013).

Definition 1 (*Convex and concave relaxation*) Let $f : \mathcal{X} \to \mathbb{R}$ with $\mathcal{X} \in \mathbb{IR}^n$. A convex function $f^{\mathrm{cv,u}} : \mathcal{X} \to \mathbb{R}$ is called a *convex relaxation* or *convex underestimator of f on $\mathcal{X}$* iff $f^{\mathrm{cv,u}}(\mathbf{x}) \leq f(\mathbf{x})$, $\forall \mathbf{x} \in \mathcal{X}$. A concave function $f^{\mathrm{cc,o}} : \mathcal{X} \to \mathbb{R}$ is called a *concave relaxation* or *concave overestimator of f* on $\mathcal{X}$ iff $f^{\mathrm{cc,o}}(\mathbf{x}) \geq f(\mathbf{x})$, $\forall \mathbf{x} \in \mathcal{X}$.

In Sects. 2.3 and 3.4, we make use of non-standard under- and overestimators to derive tighter relaxations.

Definition 2 (*Concave underestimator and convex overestimator*) Let $f : \mathcal{X} \to \mathbb{R}$ with $\mathcal{X} \in \mathbb{IR}^n$. A convex function $f^{\mathrm{cv,o}} : \mathcal{X} \to \mathbb{R}$ is called a *convex overestimator of f on $\mathcal{X}$* iff $f^{\mathrm{cv,o}}(\mathbf{x}) \geq f(\mathbf{x})$, $\forall \mathbf{x} \in \mathcal{X}$. A concave function $f^{\mathrm{cc,u}} : \mathcal{X} \to \mathbb{R}$ is called a *concave underestimator of f* on $\mathcal{X}$ iff $f^{\mathrm{cc,u}}(\mathbf{x}) \leq f(\mathbf{x})$ $\forall \mathbf{x} \in \mathcal{X}$.

[1] MAiNGO is available at https://git.rwth-aachen.de/avt.svt/public/maingo.git along with a version of MC++ implementing the proposed relaxations.

The original McCormick technique (McCormick 1976) provides rules for computing relaxations of so-called *factorable functions* consisting of finite compositions of binary sums, binary products, and a library of univariate functions. The latter are called *intrinsic functions*, and for these functions both convex and concave relaxations and range bounds need to be known.

Definition 3 (*Range bounds*) Let $f : \mathcal{X} \to \mathbb{R}$ with $\mathcal{X} \in \mathbb{IR}^n$. Scalars z^{L} and z^{U} are called *range bounds* of f over $\mathcal{X}$ iff $f(\mathcal{X}) \subseteq [z^{\mathrm{L}}, z^{\mathrm{U}}]$.

Tsoukalas and Mitsos (2014) extended the McCormick technique to allow for multivariate intrinsic functions. Note that such intrinsic functions are also used in the context of the AVM, such that the relaxations and range bounds developed herein could also be applied in that context. Unlike the AVM (Smith and Pantelides 1997; Tawarmalani and Sahinidis 2002), the McCormick technique provides convex and concave relaxations in the original variable space. A challenge in the application of the multivariate McCormick method is that the evaluation of the relaxation requires the solution of a convex but possibly nonlinear and nonsmooth problem (Theorem 2, Tsoukalas and Mitsos (2014)). To this end, it is highly desirable for the developed relaxations to have specific monotonicity properties that allow us to either determine closed-form solutions of the problem described by Theorem 2 of Tsoukalas and Mitsos (2014) analytically, or compute it numerically with relatively little effort, e.g., by solving a one-dimensional nonlinear equation, which is solved via Newton's method.

2.2 Componentwise convex functions

Componentwise convexity of multivariate functions enables the construction of tight relaxations.

Definition 4 (*Componentwise convexity*) A function $f : \mathcal{X} \to \mathbb{R}$, where $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_n \in \mathbb{IR}^n$, is said to be *componentwise convex (concave) with respect to a variable* $x_i, i \in \{1, \dots, n\}$ iff the univariate function

$$\hat{f} : \mathcal{X}_i \to \mathbb{R},\ x_i \mapsto f(\hat{x}_1, \dots, \hat{x}_{i-1}, x_i, \hat{x}_{i+1}, \dots \hat{x}_n) \tag{1}$$

is convex (concave) for any fixed $\hat{x}_j \in \mathcal{X}_j, j \neq i$. f is said to be *componentwise convex (concave)* iff it is componentwise convex (concave) with respect to each variable $x_i, i \in \{1, \dots, n\}$.

For twice continuously differentiable functions, a convenient way of confirming componentwise convexity or concavity with respect to a variable consists in examining the second derivative of the univariate function $\hat{f}$ in (1) (and hence the second partial derivative with respect to the variable of interest) and proving that it is non-negative or non-positive over the considered set (cf. Rockafellar 1970), either analytically or through global maximization or minimization. However, since some of the functions considered herein are only piecewise continuously differentiable, we use a specific

procedure that enables an analysis for functions where a partial derivative need not exist at all points. The functions we consider herein are of the form

$$f : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathbb{R},\ (x_1, x_2) \mapsto \begin{cases} f_1(x_1, x_2), & \text{if } x_2 \geq \tilde{x}_2(x_1) \\ f_2(x_1, x_2), & \text{otherwise} \end{cases}, \tag{2}$$

where $\mathcal{X}_1, \mathcal{X}_2 \in \mathbb{IR}$, f_1,f_2 are both smooth on $\mathcal{X}_1 \times \mathcal{X}_2$, f is continuous on $\mathcal{X}_1 \times \mathcal{X}_2$, and $\tilde{x}_2(x_1)$ is a strictly monotonic function. To determine whether functions of form (2) are componentwise convex, we proceed as described in the following. First, we solve the optimization problems

$$\begin{aligned} \min_{x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2} \quad & \frac{\partial^2 f_1}{\partial x_2^2} \\ \text{s.t.} \quad & x_2 \geq \tilde{x}_2(x_1), \end{aligned} \tag{3}$$

$$\begin{aligned} \min_{x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2} \quad & \frac{\partial^2 f_2}{\partial x_2^2} \\ \text{s.t.} \quad & x_2 \leq \tilde{x}_2(x_1), \end{aligned} \tag{4}$$

to global optimality. Note that these problems are much smaller and less complex than the design problems within which the functions are intended to be used and can be solved with general purpose methods. Furthermore, since we only consider certain known functions, the problems need to be solved only once in order to construct tighter relaxations for later use in design problems. If the solution values of both problems (3) and (4) are non-negative, we know that f_1 is componentwise convex with respect to x_2 on $\mathcal{S}_1 := \{(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2 | x_2 \geq \tilde{x}_2(x_1)\}$ and f_2 is componentwise convex with respect to x_2 on $\mathcal{S}_2 := \{(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2 | x_2 \leq \tilde{x}_2(x_1)\}$. To show that f is componentwise convex with respect to x_2 on $\mathcal{X}_1 \times \mathcal{X}_2$, it remains to show that the difference between the right and left derivatives of f with respect to x_2 at the boundary between the subdomains $\mathcal{S}_1$ and $\mathcal{S}_2$ is non-negative. To this end, we solve a third optimization problem given as

$$\min_{x_1 \in \mathcal{X}_1} \left. \frac{\partial f_1}{\partial x_2} \right|_{(x_1, \tilde{x}_2(x_1))} - \left. \frac{\partial f_2}{\partial x_2} \right|_{(x_1, \tilde{x}_2(x_1))} \tag{5}$$

and examine the sign of the optimal objective value. When applying the above procedure for testing componentwise convexity with respect to x_1, the order of the derivatives of f_1 and f_2 in (5) needs to be exchanged in case $\tilde{x}_2(x_1)$ is increasing. An analogous procedure can be used for testing for componentwise concavity.

Once it has been established that a function is componentwise convex (concave), its concave (convex) envelope is known to be vertex polyhedral (Tardella 2004). Since the multivariate functions considered herein are only bivariate, this envelope can be computed efficiently using the method of Meyer and Floudas (2005). A convex (concave) relaxation, on the other hand, can be computed using the method of Najman et al. (2019a) if the derivatives of the function fulfill

certain additional requirements. However, since some of the functions herein are nonsmooth, we need to adapt the latter method. First, we define partial subderivatives in analogy to the definition of subgradients for convex functions (cf. Rockafellar 1970).

Definition 5 (*Subgradient and subdifferential*) For a convex and concave function $f^{\mathrm{cv}}, f^{\mathrm{cc}} : \mathcal{X} \to \mathbb{R}$, where $\mathcal{X} \in \mathbb{IR}^n$, we call $\mathbf{s}^{\mathrm{cv}}, \mathbf{s}^{\mathrm{cc}} \in \mathbb{R}^n$ a *convex* and a *concave subgradient of* f^{cv}, f^{cc} *at* $\hat{\mathbf{x}}$, respectively, iff

$$f^{\mathrm{cv}}(\mathbf{x}) \geq f^{\mathrm{cv}}(\hat{\mathbf{x}}) + (\mathbf{s}^{\mathrm{cv}})^T(\mathbf{x} - \hat{\mathbf{x}}), \ \forall \mathbf{x} \in \mathcal{X},$$
$$f^{\mathrm{cc}}(\mathbf{x}) \leq f^{\mathrm{cc}}(\hat{\mathbf{x}}) + (\mathbf{s}^{\mathrm{cc}})^T(\mathbf{x} - \hat{\mathbf{x}}), \ \forall \mathbf{x} \in \mathcal{X}.$$

The *convex* and *concave subdifferential of* f *at* $\hat{\mathbf{x}}$ denoted by $\partial^{\mathrm{cv}} f(\hat{\mathbf{x}}), \partial^{\mathrm{cc}} f(\hat{\mathbf{x}})$, respectively, are the sets of all convex and concave subgradients of f at $\hat{\mathbf{x}}$, respectively.

Definition 6 (*Partial subderivative and subdifferential*) Let $f : \mathcal{X} \to \mathbb{R}$, where $\mathcal{X} \in \mathbb{IR}^n$, be componentwise convex (concave) with respect to some $x_i, i \in \{1, \ldots, n\}$. We call $s \in \mathbb{R}$ a *partial subderivative of* f *with respect to* x_i *at* $\hat{\mathbf{x}} \in \mathcal{X}$ iff it is a subgradient of the univariate function $\hat{f}$ as defined in (1) at $\hat{x}_i$. The set of all partial subderivatives of f with respect to x_i at $\hat{\mathbf{x}} \in \mathcal{X}$ is called the *partial subdifferential of* f *with respect to* x_i *at* $\hat{\mathbf{x}} \in \mathcal{X}$ and denoted by $\partial_i f(\hat{\mathbf{x}})$.

Using this definition of the partial subdifferential, we introduce the following adapted version of Theorem 1 from Najman et al. (2019a):

Theorem 1 *Let* $f : \mathcal{X} \to \mathbb{R}$ *be a componentwise convex function with* $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 = \left[x_1^L, x_1^U\right] \times \left[x_2^L, x_2^U\right] \in \mathbb{IR}^2$, *and* $\mathbf{x}^b$ *a border point of* $\mathcal{X}$ *with* $x_1^b = x_1^L$ *and* $x_2^b \in \left[x_2^L, x_2^U\right]$. *Define the function*

$$f_{sum}^{cv,u}(\mathbf{x}) := f\left(x_1, x_2^b\right) + f\left(x_1^b, x_2\right) - f\left(x_1^b, x_2^b\right). \tag{6}$$

If it holds that

$$\min \partial_1 f\left(x_1, \bar{x}_2\right) \geq \max \partial_1 f\left(x_1, x_2^b\right) \quad \forall x_1 \in \mathcal{X}_1 \tag{7}$$

for any fixed $\bar{x}_2 \in \mathcal{X}_2$, *then* $f_{sum}^{cv,u}$ *is a convex relaxation of* f.

Proof The proof is analogous to that of Theorem 1 of Najman et al. (2019a) when replacing the partial derivatives therein with the suitable minimum and maximum elements of the partial subdifferentials as in (7), and replacing the derivatives in Lemma 1 of Najman et al. (2019a) with the subgradients for convex (rather than differentiable) functions. □

In particular, Theorem 1 can be applied to componentwise convex functions of form (2). Similar to Najman et al. (2019a), up to 2 valid relaxations can be obtained from Theorem 1 by reordering the variables and changing the sign of the

coordinates. Furthermore, an analogous way for constructing concave relaxations of componentwise concave functions is given by considering $-f$ instead.

In Corollary 1 of Najman et al. (2019a), a convenient procedure based on mixed second-order partial derivatives is provided for confirming that the conditions of Theorem 1 therein are satisfied. This procedure is not directly applicable for confirming the validity of (7) because functions of form (2) are possibly not twice differentiable, despite the fact that f_1 and f_2 are both twice differentiable. Therefore, we describe an alternative way of confirming the validity of (7) for functions considered of form (2), which is analogous to the test for componentwise convexity described above: First, we solve the two optimization problems

$$\begin{aligned} \min_{x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2} \quad & \frac{\partial^2 f_1}{\partial x_1 \partial x_2} \\ \text{s.t.} \quad & x_2 \geq \tilde{x}_2(x_1), \end{aligned} \tag{8}$$

$$\begin{aligned} \min_{x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2} \quad & \frac{\partial^2 f_2}{\partial x_1 \partial x_2} \\ \text{s.t.} \quad & x_2 \leq \tilde{x}_2(x_1), \end{aligned} \tag{9}$$

to global optimality. If the solution values of both problems (8) and (9) are non-negative, Corollary 1 of Najman et al. (2019a) suggests that the corner $\mathbf{x}^c = (x_1^L, x_2^L)$ can be used in Theorem 1 in order to construct the convex relaxation. Assume w.l.o.g. that f equals f_2 at $\mathbf{x}^c$. In order to guarantee that the relaxation resulting at $\mathbf{x}^c$ is also valid for f, a sufficient condition is that

$$\left.\frac{\partial f_1}{\partial x_2}\right|_{(x_1, \tilde{x}_2(x_1))} - \left.\frac{\partial f_2}{\partial x_2}\right|_{(x_1, \tilde{x}_2(x_1))}$$

is non-negative for all $x_1 \in \mathcal{X}_1$ if $\tilde{x}_2(x_1)$ is decreasing and non-positive if $\tilde{x}_2(x_1)$ is increasing. Since this condition is equivalent to the one for confirming componentwise convexity via problem (5) or concavity via the equivalent maximization problem, respectively, the condition is automatically satisfied for componentwise concave functions in case $\tilde{x}_2(x_1)$ is increasing and componentwise convex functions in case $\tilde{x}_2(x_1)$ is decreasing. An analogous procedure can be used for testing whether the corner point $\mathbf{x}^c = (x_1^L, x_2^U)$ can be used in Theorem 1.

2.3 Variants of αBB

The αBB method was introduced as a general-purpose method for constructing relaxations of twice continuously differentiable functions (Androulakis et al. 1995). The basic idea is to add a quadratic function with parameters $\boldsymbol{\alpha}$ to construct convex underestimators, where the $\boldsymbol{\alpha}$ parameters are chosen such that the quadratic function offsets the smallest eigenvalue of the Hessian of the function over the considered domain. These $\boldsymbol{\alpha}$ parameters can either be precomputed and thus be independent of

the interval, or they can be computed using interval methods when constructing the relaxation for a specific interval (Adjiman et al. 1998).

Hasan (2018) recently introduced a variant of αBB that does not use the $\boldsymbol{\alpha}$ parameters to construct convex underestimators directly, but rather to construct a componentwise concave underestimator, and then uses the vertex polyhedral convex envelope of the latter (cf. Sect. 2.2) as convex underestimator. Herein, we use a similar approach in the sense that we only aim at achieving componentwise convexity or concavity and not convexity or concavity of the function directly.

Lemma 1 *Let* $f : \mathcal{X} \to \mathbb{R}$, *where* $[x_1^L, x_1^U] \times \cdots \times [x_n^L, x_n^U] = \mathcal{X} \subset \mathbb{IR}^n$, *be such that for some* $i \in \{1, \ldots, n\}$, $\frac{\partial^2 f}{\partial x_i^2}$ *exists for all* $\mathbf{x} \in \mathcal{X}$. *For any* $\alpha_i^{cv} \geq \max\left\{0, -\frac{1}{2}\min_{\mathbf{x}\in\mathcal{X}} \frac{\partial^2 f}{\partial x_i^2}\right\}$, *define*

$$f^{cv,u,\alpha BB}(\mathbf{x}) := f(\mathbf{x}) + \alpha_i^{cv}\left(x_i - x_i^L\right)\left(x_i - x_i^U\right), \tag{10}$$

$$f^{cv,o,\alpha BB}(\mathbf{x}) := f(\mathbf{x}) + \alpha_i^{cv}\left(x_i - \frac{x_i^L + x_i^U}{2}\right)^2. \tag{11}$$

Then $f^{cv,u,\alpha BB}(\mathbf{x}) \leq f(\mathbf{x}) \leq f^{cv,o,\alpha BB}(\mathbf{x})\ \forall \mathbf{x} \in \mathcal{X}$, *and* $f^{cv,u,\alpha BB}$ *and* $f^{cv,o,\alpha BB}$ *are componentwise convex with respect to* x_i.

Proof The result for $f^{cv,u,\alpha BB}$ follows when applying the original αBB method (Androulakis et al. 1995) to the univariate function $\hat{f}$ in (1). For $f^{cv,o,\alpha BB}$, it follows from the results of Hasan (2018) applied to $-f$. □

Lemma 2 *Let* f *and* $\mathcal{X}$ *be as in Lemma 1. For any* $\alpha_i^{cc} \geq \max\left\{0, \frac{1}{2}\max_{\mathbf{x}\in\mathcal{X}} \frac{\partial^2 f}{\partial x_i^2}\right\}$, *define*

$$f^{cc,u,\alpha BB}(\mathbf{x}) := f(\mathbf{x}) - \alpha_i^{cc}\left(x_i - \frac{x_i^L + x_i^U}{2}\right)^2, \tag{12}$$

$$f^{cc,o,\alpha BB}(\mathbf{x}) := f(\mathbf{x}) - \alpha_i^{cc}\left(x_i - x_i^L\right)\left(x_i - x_i^U\right). \tag{13}$$

Then $f^{cc,u,\alpha BB}(\mathbf{x}) \leq f(\mathbf{x}) \leq f^{cc,o,\alpha BB}(\mathbf{x})\ \forall \mathbf{x} \in \mathcal{X}$, *and* $f^{cc,u,\alpha BB}$ *and* $f^{cc,o,\alpha BB}$ *are componentwise concave with respect to* x_i.

Proof The proof is analogous to that of Lemma 1. □

Since both the original αBB method and the variant of Hasan (2018) require the function to be twice continuously differentiable, we consider the following procedure for two dimensional functions of form (2): If the solution of problem (5) is nonnegative, we use the minimum of the solution values of problems (3) and (4) to construct under- and overestimators as in Lemma 1 that are componentwise convex with

respect to x_1. Instead, if the solution value of the maximization problem analogous to (5) is non-positive, we solve the maximization problems analogous to problems (3) and (4) and use the maximum of their solution values to construct under- and overestimators as in Lemma 2 that are componentwise concave with respect to x_1. The same procedure is used to construct under- and overestimators that are componentwise convex or concave with respect to x_2.

In the present work, we only use αBB to compute relaxations of intrinsic functions within larger factorable functions, which are then relaxed using the multivariate McCormick theorem (cf. Sect. 2.1). Therefore, we need to be able to compute minima of convex relaxations and maxima of concave relaxations, respectively, over boxes, either analytically or with little effort. However, the αBB-type relaxations (10)–(13) may not be monotonic even when the original functions are, thus complicating the application of the multivariate McCormick theorem. We therefore add a linear term to functions described in (10)–(13) to make the final αBB relaxation monotonic. While this makes the final relaxations slightly weaker, it greatly simplifies the application of the multivariate McCormick theorem.

In order to apply Theorem 1, it may additionally be necessary to alter the mixed second-order partial derivatives of a function f, which we achieve with *mixed* αBB terms. For example, to obtain an underestimator that has a non-negative mixed second-order partial derivative with respect to x_i and x_j, where $i, j \in \{1, \dots, n\}$, throughout its domain, we use

$$f^{pos,u,\alpha BB}(\mathbf{x}) := f(\mathbf{x}) + \alpha_{i,j}^{\text{pos}}(x_i - x_i^L)(x_j - x_j^U)$$

with $\alpha_{i,j}^{\text{pos}} \geq \max\left\{0, -\frac{1}{2}\min_{\mathbf{x}\in\mathcal{X}} \frac{\partial^2 f}{\partial x_i \partial x_j}\right\}$ in analogy to Lemma 1. This procedure can be used analogously to achieve a non-positive mixed second-order partial derivative and for overestimators.

3 Construction of relaxations of functions from the IAPWS-IF97

The IAPWS-IF97 is divided into five regions (cf. Fig. 1). Region 1 represents subcooled liquid up to a temperature of $T_1^{\max} = 623.15\,\text{K}$, Region 2 superheated vapor, and Region 4 the two-phase region. Of the latter, we only consider the part up to $T_3^{\min} = T_1^{\max}$ (and, accordingly, $p_3^{\min} \approx 16.5292\,\text{MPa}$) in which the saturated liquid and vapor states lie in Regions 1 and 2, respectively. This part will be denoted as *Region 4-1/2* in the following. Additionally, we restrict Region 2 to pressures $p \geq p_2^{\min} := p_1^{\min} = 611.2127\,\text{Pa}$ instead of the original definition $p \geq 0$ (IAPWS 2007a) to avoid the difficulty that entropy goes to infinity as pressure goes to zero. Since such low pressures are not typically encountered in steam cycles, this restriction of Region 2 is not a practical limitation for the intended application. We do not consider Region 5, since it corresponds to much higher temperatures than those encountered in steam cycles. Region 3, which represents

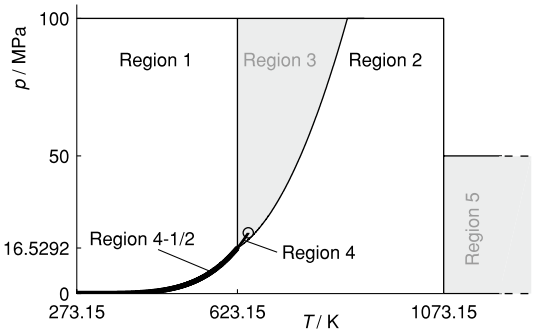

Fig. 1 Regions of the revised release of the IAPWS-IF97 (IAPWS 2007a). Only Regions 1, 2, and 4-1/2 are considered in this work, where Region 4-1/2 is the part of Region 4 adjacent to Regions 1 and 2

the region around and above the critical point, can be relevant for steam cycles and could be included as future work.

Within the considered regions, we are interested in those functions that relate temperature T, pressure p, specific enthalpy h, specific entropy s, and vapor fraction y. Specifically, we consider the following functions taken directly from the IAPWS-IF97:

- Region 1: $h_1(p,T)$, $s_1(p,T)$, $T_1(p,h)$, $T_1(p,s)$
- Region 2: $h_2(p,T)$, $s_2(p,T)$
- Region 4-1/2: $p_s(T)$, $T_s(p)$
- Boundary between Regions 2 and 3: $p_{B23}(T)$, $T_{B23}(p)$

In Region 2, we do not consider the functions $T_2(p,h)$ and $T_2(p,s)$ from the IAPWS-IF97 although they would be useful to eliminate optimization variables in reduced-space optimization formulations (Bongartz and Mitsos 2017). The reason is that these functions have piecewise definitions over subdomains in the p–h or p–s space. Since the functions were independently fitted to the original data for each of these subdomains during the development of the IAPWS-IF97, the resulting functions $T_2(p,h)$ and $T_2(p,s)$ are not continuous over their entire domains. Although the differences between the one-sided limits at the interface between different subdomains is within the tolerated error for model development (Wagner et al. 2000), this discontinuity does complicate the analysis of the functions for deriving relaxations. Similar complications arise in Region 3, for which useful supplementary functions with piecewise definitions have been introduced (IAPWS 2007b).

In addition to the functions taken directly from the IAPWS-IF97, we define the following auxiliary functions for calculating specific enthalpy and entropy of saturated vapor and liquid states:

$$h^{\mathrm{liq}}_{4-1/2}(p) := h_1(p, T_s(p)), \tag{14}$$

$$h^{\text{vap}}_{4-1/2}(p) := h_2(p, T_s(p)), \tag{15}$$

$$s^{\text{liq}}_{4-1/2}(p) := s_1(p, T_s(p)), \tag{16}$$

$$s^{\text{vap}}_{4-1/2}(p) := s_2(p, T_s(p)). \tag{17}$$

Using (14)–(17), we define the following functions for computing specific enthalpy and entropy as well as vapor fraction in the two-phase region:

$$h_{4-1/2}(p, y) := y h^{\text{vap}}_{4-1/2}(p) + (1-y) h^{\text{liq}}_{4-1/2}(p), \tag{18}$$

$$s_{4-1/2}(p, y) := y s^{\text{vap}}_{4-1/2}(p) + (1-y) s^{\text{liq}}_{4-1/2}(p), \tag{19}$$

$$y_{4-1/2}(p, h) := \frac{h - h^{\text{liq}}_{4-1/2}(p)}{h^{\text{vap}}_{4-1/2}(p) - h^{\text{liq}}_{4-1/2}(p)}, \tag{20}$$

$$y_{4-1/2}(p, s) := \frac{s - s^{\text{liq}}_{4-1/2}(p)}{s^{\text{vap}}_{4-1/2}(p) - s^{\text{liq}}_{4-1/2}(p)}. \tag{21}$$

For the bivariate functions, we define two types of domains that are useful for the analysis of the functions.

Definition 7 (*Physical domain*) Given a function $f(x_1, x_2)$, we denote by $\mathcal{P}_{f(x_1,x_2)} \subset \mathbb{R}^2$ the domain specified in the IAPWS-IF97 (IAPWS 2007a) and we call $\mathcal{P}_{f(x_1,x_2)}$ the *physical domain* of $f(x_1, x_2)$.

Definition 8 (*Box domain*) Given a function $f(x_1, x_2)$ with physical domain $\mathcal{P}_{f(x_1,x_2)} \subset \mathbb{R}^2$, we denote by $\mathcal{B}_{f(x_1,x_2)} \in \mathbb{IR}^2$ the smallest box containing $\mathcal{P}_{f(x_1,x_2)}$ and call $\mathcal{B}_{f(x_1,x_2)}$ the *box domain* of $f(x_1, x_2)$.

In the following, the lower and upper variable bounds of the box domains are denoted with the superscripts *min* and *max*, respectively, i.e., $\mathcal{B}_{f(x_1,x_2)} = [x_1^{\min}, x_1^{\max}] \times [x_2^{\min}, x_2^{\max}]$. The physical domains, which are often nonconvex (cf. Fig. 1), can be represented in terms of the box domains and additional inequalities. This representation is useful because in global optimization the functions typically need to be evaluated on rectangular subsets of the box domains (called *nodes* of the B&B tree) while the inequalities are enforced as constraints. To distinguish the bounds of nodes from those of the box domains, we denote the former by the superscripts L and U, e.g., $[x_1^L, x_1^U] \times [x_2^L, x_2^U] \subseteq \mathcal{B}_{f(x_1,x_2)}$.

All functions of the IAPWS-IF97 and by extension also those in (14)–(21) are given as explicit expressions, in most cases compositions of linear and signomial functions, and can be written as factorable functions. Therefore, general purpose methods for constructing relaxations of factorable functions such as the (multivariate) McCormick technique (McCormick 1976; Tsoukalas and Mitsos 2014) or the auxiliary variable method (Smith and Pantelides 1997; Tawarmalani and Sahinidis 2002) could be applied to the original, factorable representation of the functions. However, this is problematic for two reasons. First, the functional forms of the expressions are rather long and complex. Therefore, general purpose methods often result in rather weak relaxations. Second, during optimization, boxes need to be considered that contain regions of the box domain that lie outside the physical domain. Even though such regions are made infeasible using a suitable constraint, the functions may still need to be evaluated at points in such regions and relaxations need to be constructed over boxes containing these regions as well. However, the functions were never intended to be used outside their physical domain, and many functions take values of extremely large magnitude when evaluated in such regions. This can lead to very weak relaxations over the physical domain when considering boxes containing these regions, as well as numerical problems when solving the linear and nonlinear subproblems.

To address these challenges, we consider the IAPWS-IF97 functions directly as intrinsic functions instead of using their factorable representations. To derive the required information to treat them as intrinsic functions, we conduct the following steps that are explained in detail in the following subsections:

1. We determine the physical domains and box domains for bivariate functions.
2. We modify the bivariate functions outside their physical domains to improve the mathematical properties of the functions in these regions.
3. We analyze the functions for monotonicity properties to derive range bounds.
4. We analyze the functions for convexity properties to derive convex and concave relaxations.

3.1 Determination of physical domains and box domains

For the bivariate functions, we first determine the physical domains and box domains according to Definitions 7 and 8. For the functions $h_1(p,T)$, $s_1(p,T)$, $h_2(p,T)$, and $s_2(p,T)$, which are obtained as derivatives of the *basic equations* of the IAPWS-IF97, the domains can be taken directly from the model definition (IAPWS 2007a). For most other functions, the physical domains are not explicitly given in the IAPWS-IF97. In this case, the physical domains need to be determined through an analysis of the monotonicity properties of related functions from the IAPWS-IF97 and the box domains need to be determined via globally maximizing and minimizing each variable over the physical domain. Examples for this procedure both for a straightforward case and a more involved case can be found in "Appendix 1" along with a summary of all box and physical domains in Tables 5 and 6.

3.2 Model modification outside the physical domain

We analyze the model functions and their derivatives to determine whether the functions exhibit excessive peaks or other undesired behavior in those regions of their box domains that are outside the physical domain. In case of undesired behavior, we replace the functions with a suitable extrapolation outside (a relaxation of) the physical domain. The resulting piecewise defined functions are of form (2) and will be called *intermediate modified functions* in the following and denoted with the superscript *int*. The extrapolations are chosen such that

- they have similar monotonicity and convexity properties as the original functions on their physical domains as far as possible,
- the intermediate modified functions are continuous,
- the solution values of problem (5) for the intermediate modified functions are nonnegative, or those of the corresponding maximization problems are nonpositive (whenever possible, both are achieved simultaneously by making the functions continuously differentiable).

Beyond these intermediate modifications, we restrict the range of each function to that achieved over its physical domain. This helps to avoid domain violations when considering compositions of functions from the IAPWS-IF97. The resulting functions will be called *final modified functions* in the following and denoted with the superscript *mod*.

The final modified functions are the ones that will be used in the optimization problems. The relaxations, however, will be constructed based on the intermediate modified functions that have more useful convexity properties (cf. Sect. 3.4). Based on these relaxations of the intermediate modified functions, relaxations for the final modified functions are obtained using the rules for composition with the max and min functions (Tsoukalas and Mitsos 2014).

Figure 2 shows the application of the above procedure to the function $h_2(p, T)$. A more detailed description for this function can be found in "Appendix 2". Similar modifications are conducted for the functions $h_1(p, T)$, $s_1(p, T)$, $T_1(p, h)$, and $s_2(p, T)$. The intermediate modifications are summarized in Table 7 in "Appendix 2". The remaining bivariate functions are only cut at the minimum and maximum values occurring on their physical domains but not extended otherwise since their properties outside their physical domains are already satisfactory.

Note that the modifications of the model functions themselves are only conducted in parts of the box domain where the model has no physical meaning and that are excluded via suitable constraints when using the functions in an optimization problem. They thus do not alter the solutions of meaningful process models using these functions. Note also that the modifications do make the functions nonsmooth outside of the physical domains (both the replacement of the function with an extrapolation for some of the intermediate modified function and the restriction of the range of the functions to that over their physical domains when constructing the final modified functions). However, we did not find this to be an issue in practice so far (cf. discussion in Sect. 4).

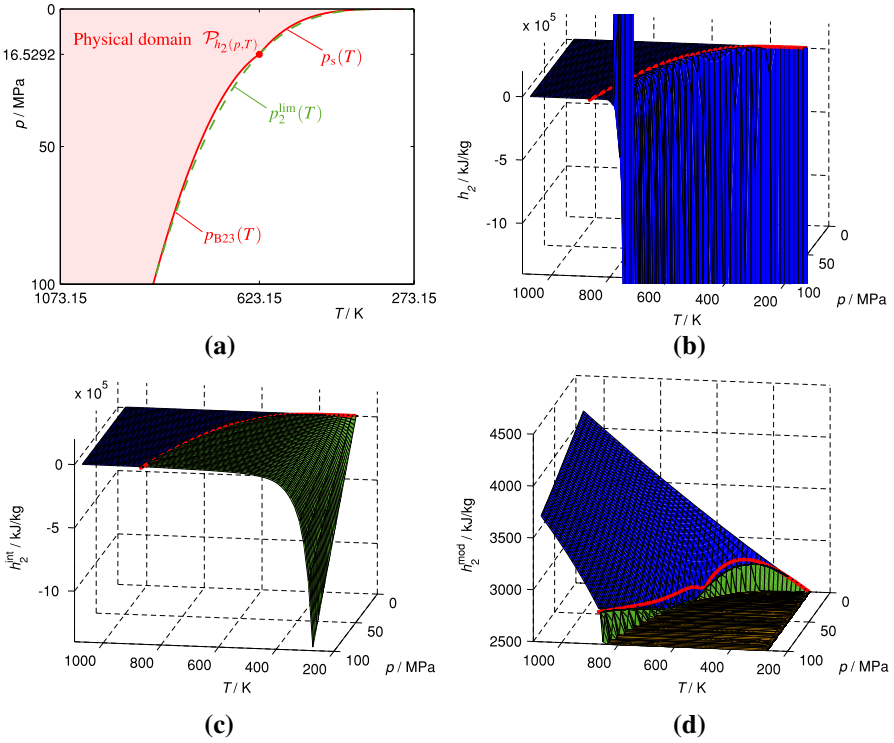

Fig. 2 Modification procedure for the function $h_2(p,T)$ as described in "Appendix 2". In all subplots, the solid red line denotes the boundary of the physical domain. **a** The physical domain of $h_2(p,T)$, $\mathcal{P}_{h_2(p,T)}$, is a nonconvex subset of the box domain $[611.2127 \times 10^{-6}, 100]\,\text{MPa} \times [273.15, 1073.15]\,\text{K}$ delimited by $p \le p_s(T)$ for $T \le 623.15\,\text{K}$ and $p \le p_{B23}(T)$ otherwise (note that both axes are inverted). Since this boundary is nonsmooth, we introduce a relaxed physical domain delimited by the smooth function $p_2^{\lim}(T)$ instead. **b** Original function $h_2(p,T)$ with undesired peaks outside $\mathcal{P}_{h_2(p,T)}$ that go beyond 4.5×10^9 kJ/kg and -6.1×10^{55} kJ/kg and were cut off to improve readability. **c** Intermediate modified function $h_2^{\text{int}}(p,T)$ that was modified (green) for $p > p_2^{\lim}(T)$ to have favorable properties for deriving relaxations. **d** Final modified function $h_2^{\text{mod}}(p,T) = \max(h_2^{\text{int}}(p,T), h_2^{\min})$, where $h_2^{\min}$ is the minimum of $h_2(p,T)$ over $\mathcal{P}_{h_2(p,T)}$. The region where this lower bound is active is shown in orange. Note the different scale of the z-axes compared with **b** and **c**

3.3 Range bounds

To derive range bounds, we analyze the intermediate modified functions for monotonicity properties by globally maximizing and minimizing their first partial derivatives over the corresponding box domains using our open-source global solver MAiNGO (Bongartz et al. 2018). For functions that are replaced by an extrapolation outside their physical domains and that are hence of form (2), the maximization and minimization is done separately within the relaxed physical domain and the domain of the extrapolation. In both subdomains, the functions are differentiable by construction. Even though the functions may be nonsmooth at the boundary between the subdomains, monotonicity results within the subdomains can still be translated

into global monotonicity properties since the functions are continuous. These properties also hold for the final modified functions since taking the maximum or minimum with a constant does not change the monotonicity properties.

For functions with monotonicity properties over the entire (box) domain, we immediately obtain exact range bounds. For other functions, exact range bounds can only be obtained in certain cases. In case a function is only monotonic with respect to one variable, an exact upper or lower bound can sometimes still be obtained in case of componentwise convexity or concavity with respect to the other variable (cf. Sect. 3.4). As a last resort, we obtain natural interval extensions from the FILIB++ library (Lerch et al. 2011) by evaluating the factorable representation of the function using the interval datatypes from FILIB++ either with respect to one or both variables over a suitable part of its domain. Examples for both a straightforward case and a more involved case can be found in "Appendix 3", and all exploited monotonicity properties are summarized in Table 8.

3.4 Convex and concave relaxations

Similar to the monotonicity analysis described above, we analyze the intermediate modified functions for (componentwise) convexity by globally maximizing and minimizing their second partial derivatives over their box domains using MAiNGO. For bivariate functions with piecewise definition of form (2), we proceed as described in Sect. 2.2. The identified properties are summarized in Table 9 in "Appendix 4". These properties are used to derive relaxations as described in the following.

3.4.1 Univariate functions

For univariate functions that are convex or concave over their entire domain (cf. Table 9), we trivially obtain the convex and concave envelopes over any interval as the function itself and the secant between the end points of the interval. The functions $h^{\text{liq}}_{\text{4-1/2}}(p)$, $s^{\text{liq}}_{\text{4-1/2}}(p)$ and $s^{\text{vap}}_{\text{4-1/2}}(p)$, on the other hand, are not convex or concave on their entire domains. However, they are either mostly convex or mostly concave, while the remaining parts are essentially linear. We can therefore obtain rather tight relaxations via the αBB variants (10)–(13) with fixed values for the α parameters. These values are obtained as described in Lemmata 1 and 2 from the maximum or minimum values (whichever is smaller in magnitude) of the second derivatives over the entire domain that were precomputed through global optimization in MAiNGO. When constructing relaxations on a given interval during the B&B procedure, the αBB terms in (10)–(13) are only used if the node is not fully in a region where the function is convex or concave (cf. Table 9). Since the functions are univariate, convex and concave envelopes could instead also be obtained using the technique described by McCormick (1976, Section 4). However, in general this method requires iterative solution of a nonlinear equation using, e.g., Newton's method, for each evaluation of the relaxations, and we would like to avoid this computational effort.

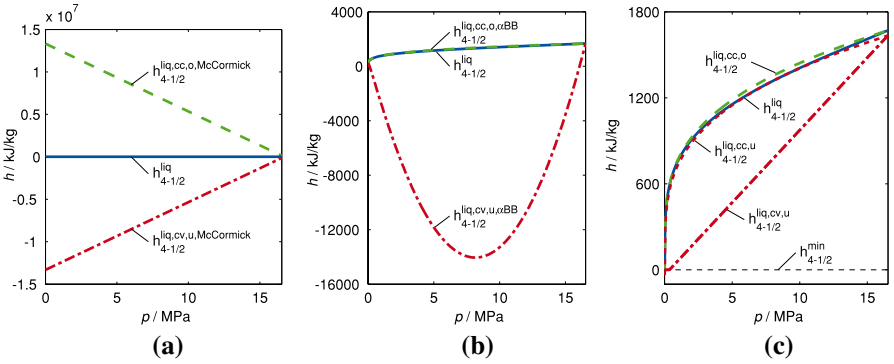

Fig. 3 Convex (cv) and concave (cc) relaxations of the function $h^{\text{liq}}_{4-1/2}(p)$ constructed using different methods. Note the different scales on the y-axis. **a** McCormick relaxations applied to $h^{\text{liq}}_{4-1/2}(p) = h_1(p, T_s(p))$, where the relaxations of $T_s(p)$ needed to be artificially bounded to positive values to avoid domain violations. **b** αBB relaxations (10) and (13) using exact α values (determined through global optimization) for the considered domain. **c** Relaxations used in the present work as described in "Appendix 4". The dotted line denotes the concave underestimator (cc,u) from the αBB variant (12). The dash-dotted convex relaxation is the secant of this concave underestimator, additionally cut at the lower bound $h^{\text{min}}_{4-1/2} = \min_{p\in[p^{\text{min}}_{4-1/2}, p^{\text{max}}_{4-1/2}]} h^{\text{liq}}_{4-1/2}(p)$

As an example, Fig. 3c shows the proposed relaxations for $h^{\text{liq}}_{4\text{-}1/2}(p)$, which are discussed in more detail in "Appendix 4". The relaxations are orders of magnitude tighter than the McCormick relaxations (cf. Fig. 3a). Quantitatively, the maximum difference between the value of $h^{\text{liq}}_{4\text{-}1/2}(p)$ and that of the convex relaxation is approximately 1.3×10^7 for the McCormick relaxation whereas it is less than 750 for the proposed relaxation. The proposed convex relaxation, which is based on the secant of the concave underestimator (12) and hence the maximum of the second derivative of $h^{\text{liq}}_{4\text{-}1/2}(p)$, is also much tighter than the regular αBB relaxation (10) based on the minimum of the second derivative of $h^{\text{liq}}_{4\text{-}1/2}(p)$. For this regular αBB relaxation, the maximum difference between the values of the function and that of the relaxation is approximately 1.5×10^4 (cf. Fig. 3b).

Beyond the tightness of the relaxations, another factor that can impact the performance of a global solver is the computational cost for computing the relaxations. To quantify this computational cost, we evaluated both the McCormick relaxations and the proposed relaxations (using the implementation described in Sect. 4) for all considered univariate functions on 100 evenly spaced points on each of 1000 randomly generated intervals within their domains and measured the CPU time. The measured time includes the time for computing the proposed range bounds in case of the proposed relaxations and that for computing natural interval extensions in case of the McCormick relaxations. For all functions, the evaluation of the proposed relaxations was 53–96% faster than that of the McCormick relaxations.

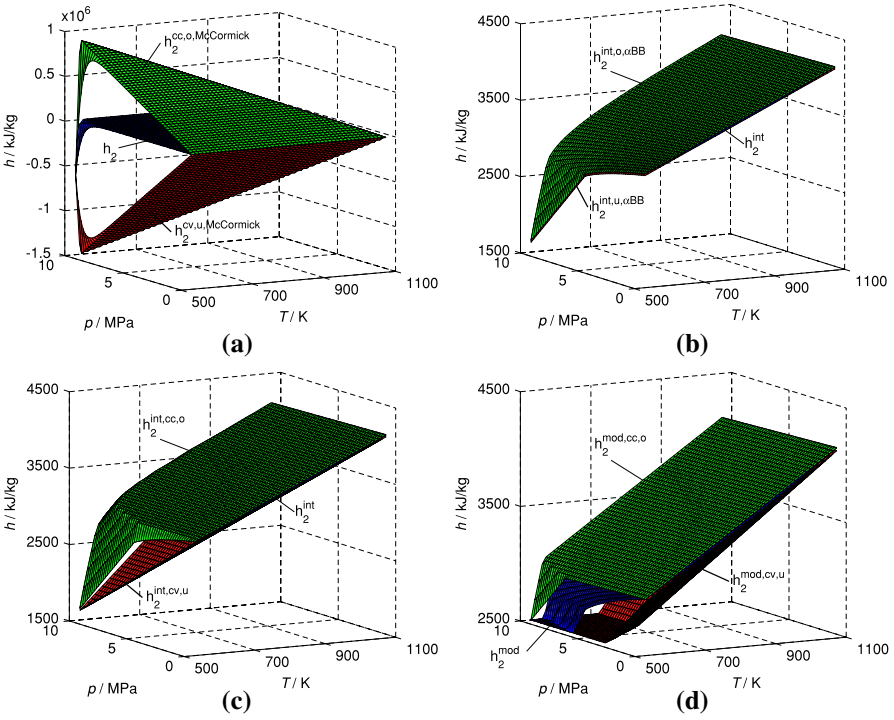

Fig. 4 Construction of relaxations of the function $h_2^{\text{mod}}(p,T)$ on $[611.2127 \times 10^{-6}, 10]\,\text{MPa} \times [520, 1073.15]\,\text{K}$. Note the different scales on the z-axes. **a** The McCormick relaxations of the original function $h_2(p,T)$ are very weak. Note also the strong negative peak of the function itself at high p and low T. **b** For the function $h_2^{\text{int}}(p,T)$, this peak is much less pronounced. The componentwise concave over- and underestimators constructed via αBB are almost identical to $h_2^{\text{int}}(p,T)$. **c** The convex and concave relaxations of $h_2^{\text{int}}(p,T)$ are constructed from these componentwise concave over- and underestimators. **d** The relaxations of the function $h_2^{\text{mod}}(p,T)$ are constructed by applying the rules for composition with the max function

3.4.2 Bivariate functions

The considered bivariate intermediate modified functions are neither convex nor concave on most parts of their domains. However, they are often componentwise convex or concave with respect to one or both variables on large parts of their domain (cf. Table 9). Additionally, they are often *almost* componentwise convex or concave over the entire domain in the sense that either the maximum of the second partial derivative is much larger in magnitude than the minimum or vice versa. This property is exploited to construct relaxations as described in the following, where we consider the function $h_2^{\text{mod}}(p,T)$ as an example.

First, in addition to the identification of regions where the intermediate modified function (e.g., $h_2^{\text{int}}(p,T)$) is componentwise convex or concave with respect to some variable, we identify the minimum and maximum values of its second partial derivatives via global optimization using MAiNGO (separately over the subdomains in

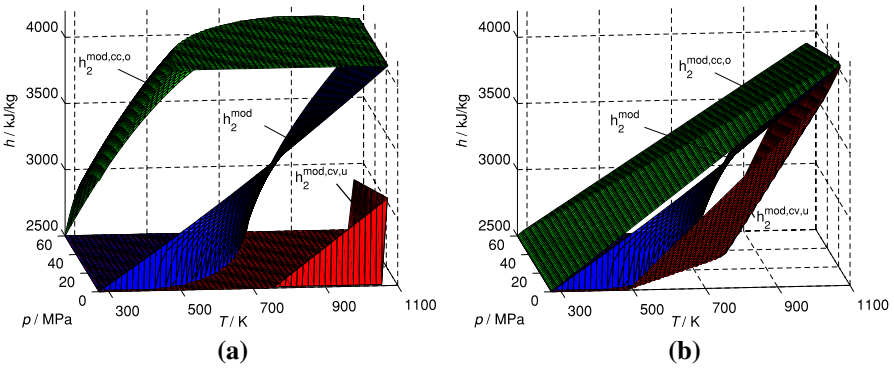

Fig. 5 Relaxations for the function $h_2^{\text{mod}}(p,T)$ on $[611.2127 \times 10^{-6}, 60]\,\text{MPa} \times [273.15, 1073.15]\,\text{K}$. **a** The relaxations constructed as in Fig. 4 are somewhat weak over larger boxes. **b** We therefore add ad hoc relaxations that do not converge but are valid over the entire domain and provide tighter relaxations on large boxes

case of piecewise defined functions, cf. Sect. 2.3). For those regions where the function is not componentwise convex with respect to a variable, we apply componentwise αBB with respect to that variable using either (10) and (11) or (12) and (13), depending on whether the function is *almost* componentwise convex or concave with respect to that variable in the above sense (cf. Fig. 4b). If necessary, we add a linear term to (10)–(13) to ensure that the relaxation is monotonic, and we include a mixed αBB term to ensure a constant sign of the mixed second-order partial derivatives (cf. Sect. 2.3).

In most cases, the resulting under- and overestimators are already either componentwise convex or concave (with respect to both variables). If, however, an underestimator is componentwise convex with respect to one variable and componentwise concave with respect to the other, we construct a componentwise convex underestimator by taking the secant with respect to the concave variable (Najman et al. 2019a). For an overestimator, we instead construct a componentwise concave overestimator by taking the secant with respect to the convex variable. For componentwise convex underestimators and componentwise concave overestimators, we use Theorem 1 to obtain convex underestimators and concave overestimators, respectively. For componentwise concave underestimators and componentwise convex overestimators, we use the method of Meyer and Floudas (2005) instead (cf. Fig. 4c).

Up to this point, the relaxations were constructed based on the intermediate modified functions (e.g., $h_2^{\text{int}}(p,T)$) before cutting at the minimum and maximum values over the physical domain. We therefore apply the rules for composition with max and min functions (Tsoukalas and Mitsos 2014) to derive valid relaxations for the final modified functions. For $h_2^{\text{mod}}(p,T)$, the relaxations are shown in Fig. 4d. They are significantly tighter than the McCormick relaxations of $h_2(p,T)$ (cf. Fig. 4a).

In some cases, however, the resulting relaxations are still somewhat weak when considering large boxes, in particular when requiring large α values for achieving componentwise convexity (cf. Fig. 5a). In these cases, we manually construct

additional (linear or nonlinear) *ad hoc relaxations*. These ad-hoc relaxations are valid over the entire box domain and are independent of the subset of the box domain over which they are evaluated in a B&B procedure. Therefore, they do not converge to the function when considering boxes of decreasing size, but they help tighten the relaxations for large boxes (cf. Fig. 5b). To construct the ad-hoc relaxations, we visually inspect the graph of the respective functions by plotting them in MATLAB. We then simultaneously plot linear or convex nonlinear functions, the potential ad-hoc relaxations, with parameters that we adjust through trial and error until the functions visually appear to be valid relaxations (e.g., in case of a convex relaxation, they appear to be below the original function on the entire box domain). We then confirm the validity of these potential ad-hoc relaxations by globally minimizing the difference between them and the original function in MAiNGO and examining the sign of the optimal objective value.

Finally, to evaluate the proposed relaxations at a given point, we select the maximum among the described convex relaxations and the minimum among the concave relaxations, including the multiple facets that are obtained from the application of both Theorem 1 and the method of Meyer and Floudas (2005). Additionally, we cut the relaxations off at the lower and upper range bounds.

The full procedure described above is used for the functions $h_1(p,T)$, $s_1(p,T)$, $T_1(p,h)$, $h_2(p,T)$, $s_2(p,T)$, and $y_{4-1/2}(p,h)$. For the functions $T_1(p,s)$, $h_{4-1/2}(p,y)$, $s_{4-1/2}(p,y)$, and $y_{4-1/2}(p,s)$, the McCormick relaxations of the factorable representations or the compositions using the relaxations of the univariate functions discussed in Sect. 3.4.1 are relatively good already so that these are used instead. For $T_1(p,s)$ and $y_{4-1/2}(p,s)$, we merely add ad-hoc relaxations for large boxes as described above.

While the proposed relaxations of the considered bivariate functions are at least as tight as and typically significantly tighter than the McCormick relaxations, unlike in the univariate case, they are not always cheaper to evaluate. When evaluating them on 100 evenly spaced points on each of 1000 randomly generated boxes within their box domains, only the proposed relaxations of $h_2(p,T)$, $s_2(p,T)$, $h_{4-1/2}(p,y)$, $s_{4-1/2}(p,y)$, and $y_{4-1/2}(p,s)$ are faster to evaluate (80–99%) than the McCormick relaxations. The time for evaluating the proposed relaxation for $T_1(p,s)$ is virtually the same as that for evaluating the McCormick relaxation, since we merely add an ad-hoc relaxation, which is very cheap to compute. For the functions $h_1(p,T)$, $s_1(p,T)$, $T_1(p,h)$, and $y_{4-1/2}(p,h)$, however, the proposed relaxations take 160–6500% longer to evaluate than the McCormick relaxations. The reason for this higher computational cost is that for these functions, iterative solution of one-dimensional nonlinear equations via Newton's method is required to determine the correct point to be used in the multivariate McCormick composition theorem (cf. Sect. 2.1). Nevertheless, computational experiments confirmed that this additional effort for computing the relaxations typically pays off in global optimization thanks to the much better tightness than the McCormick relaxations.

4 Case studies

To demonstrate the benefit of the derived relaxations, we use the design problems from our previous work (Bongartz and Mitsos 2017) that consider bottoming cycles for combined power plants in three levels of complexity:

- *Case Study 1* Basic Rankine cycle with fixed temperature of the hot gas turbine exhaust at the outlet of the heat recovery steam generator (HRSG).
- *Case Study 2* Basic Rankine cycle but with variable gas outlet temperature and with a turbine bleed to an open feedwater heater that also serves as deaerator.
- *Case Study 3* Dual-pressure cycle where the outlet of the high-pressure (HP) turbine is mixed with the outlet of the low-pressure (LP) superheater before entering the LP turbine. The latter also has a bleed stream to the deaerator.

Two objective functions are considered for each flowsheet, namely maximization of the net power output ($\dot{W}_{\text{net}}$) and minimization of the levelized cost of electricity (LCOE). The latter objective is more complex because it includes investment cost and hence, the models contain sizing and cost correlations for the process units.

In our previous work (Bongartz and Mitsos 2017), we demonstrated the benefit of modeling the design problems in a *reduced-space* optimization formulation where only the design variables and very few model variables remain in the optimization problem and most other model variables (and equations) are collapsed in a sequential evaluation of the model going through the cycle. In this previous work, we used very simple models for the thermodynamic properties of water, mostly ideal gas and liquid with constant heat capacities and constant heat of vaporization at a reference temperature. In the present work, we consider the same design problems and replace the simple thermodynamic models with the IAPWS-IF97 and solve the resulting problems using MAiNGO.

4.1 Modeling and implementation

Compared to the original reduced-space formulation (Bongartz and Mitsos 2017), we need to make some slight changes to the model. These changes are required because the functions $T_2(p,h)$ and $T_2(p,s)$ for computing temperatures for given pressure and enthalpy or entropy in the gas phase are not available (cf. discussion in Sect. 3). Specifically, we need to use steam temperatures at the superheater outlets as optimization variables, either instead of the corresponding enthalpies as degrees of freedom (for Case Studies 2 and 3) or as an additional variable with a corresponding equality constraint (for Case Study 1). For Case Study 3, we also need to add the steam temperature at the inlet of the LP turbine, i.e., after mixing the HP turbine outlet with the LP superheater outlet, as well as the temperature of the hypothetical isentropic state at the outlet of the HP turbine. Additionally, we use mass flow rates of all branches of the flowsheet as design variables instead of the overall mass flow rate and split fractions, which we found to improve the relaxations.

Table 1 Ranges and optimal values for optimization variables and objectives of Case Study 1 (basic Rankine cycle) with the IAPWS-IF97 and ideal models (in parentheses)

Symbol	Description	Unit	Range	max $\dot{W}_{net}$	min LCOE
Optimization variables					
p_{S2}	Upper cycle pressure	MPa	[0.3, 10]	10 (5.46)	4.33 (3.55)
$\dot{m}$	Cycle mass flow rate	kg/s	[5, 100]	27.9 (29.5)	28.9 (30.6)
T_{S5}	Live steam temp.	K	[300, 873]	826 (668)	751 (616)
Objective functions					
$\dot{W}_{net}$	Net power output	MW		31.7 (30.0)	
LCOE	Levelized cost of el.	\$/MWh			50.8 (50.2)

el., electricity; temp., temperature

We implement the models in C++ using two different ways of handling the IAPWS-IF97 functions: using their factorable representation, and considering them as intrinsic functions. In the former case, the factorable representation of each IAPWS-IF97 function is incorporated into the directed acyclic graph (DAG) built in MAiNGO by the MC++ library (Chachuat et al. 2015). Range bounds are then obtained via natural interval extensions by FILIB++ (Lerch et al. 2011), relaxations and their subgradients via the multivariate McCormick technique by MC++ , and gradients via automatic differentiation by FADBAD++ (Bendtsen and Stauning 2012). In the latter case, each IAPWS-IF97 function occurs as a single node in the DAG. In this case, the range bounds and relaxations derived in the previous sections are used instead. For gradients, we use automatic differentiation via FADBAD++ for each of the piecewise defined regions of the functions and arbitrarily use one of the one-sided limits at the nonsmooth kinks (the same is done for the max and min functions). While we are aware that this could potentially cause difficulties for the local subsolvers that rely on gradients, we did not experience such difficulties in the case studies (cf. Sect. 4.2). This is likely due to the fact that the performance of the local subsolvers for upper bounding is often not as crucial for the performance of global solvers. Nevertheless, in the future, it would be desirable to use recently published methods for generalized derivatives (Khan and Barton 2015) instead. Finally, we use custom relaxations for the equipment cost correlations and pressure factors for heat exchangers and deaerator vessels (Najman et al. 2019b), and for the logarithmic mean temperature difference in heat exchangers (Mistry and Misener 2016; Najman and Mitsos 2016). The process models are available via our website,[2] while the functions from the IAPWS-IF97 and the proposed relaxations are available in the MC++ library used by MAiNGO.

All problems are solved with MAiNGO v0.2.0 (Bongartz et al. 2018) using CLP 1.17.0 (Forrest et al. 2019) for the linear lower bounding problems constructed on the basis of the subgradients of the relaxations, Ipopt 3.12.12 (Wächter and Biegler

[2] The C++ implementation of the process models is available at http://permalink.avt.rwth-aachen.de/?id=409863.

Table 2 Ranges and optimal values of optimization variables and objectives of Case Study 2 (basic Rankine cycle with bleed) with the IAPWS-IF97 and ideal models (in parentheses)

Symbol	Description	Unit	Range	max $\dot{W}_{\text{net}}$	min LCOE
Optimization variables					
p_{S2}	Deaerator pressure	MPa	[0.02, 0.5]	0.02 (0.02)	0.0404 (0.02)
p_{S4}	Upper cycle pressure	MPa	[0.3, 10]	10 (4.53)	6.67 (4.03)
$\dot{m}_{Cond}$	Mass flow rate condenser	kg/s	[1, 99]	26.7 (24.5)	24.4 (25.6)
$\dot{m}_{Bleed}$	Mass flow rate bleed	kg/s	[0.05, 20]	1.45 (0.83)	2.09 (0.92)
T_{S7}	Live steam temp.	K	[300, 873]	817 (873)	758 (731)
Objective functions					
$\dot{W}_{net}$	Net power output	MW		35.7 (34.4)	
LCOE	Levelized cost of el.	\$/MWh			51.2 (48.8)

el., electricity; temp., temperature

Table 3 Ranges and optimal values of optimization variables and objectives of Case Study 3 (dual-pressure cycle) with the IAPWS-IF97 and ideal models (in parentheses)

Symbol	Description	Unit	Range	max $\dot{W}_{\text{net}}$	min LCOE
Optimization variables					
p_{S2}	Deaerator pressure	MPa	[0.02, 0.5]	0.02 (0.02)	0.04 (0.02)
p_{S4}	LP pressure level	MPa	[0.3, 1.5]	0.589 (0.920)	1.5 (1.5)
p_{S8}	HP pressure level	MPa	[1, 10]	10 (10)	6.31 (4.58)
$\dot{m}_{\text{HP}}$	Mass flow rate HP part	kg/s	[2.5, 95]	26.4 (22.5)	24.3 (21.3)
$\dot{m}_{\text{Cond}}$	Mass flow rate condenser	kg/s	[4, 99]	29.6 (28.4)	25.8 (27.1)
$\dot{m}_{\text{Bleed}}$	Mass flow rate bleed	kg/s	[0.05, 20]	1.56 (1.02)	2.25 (0.98)
T_{S7}	LP steam temp.	K	[300, 873]	584 (585)	478 (515)
T_{S11}	HP steam temp.	K	[300, 873]	873 (873)	784 (797)
T_{S12s}	Isentropic temp. HP turbine	K	[300, 873]	458 (513)	565 (623)
T_{S13}	Temp. LP turbine inlet	K	[300, 873]	508 (558)	569 (610)
Objective functions					
$\dot{W}_{\text{net}}$	Net power output	MW		39.8 (39.3)	
LCOE	Levelized cost of el.	\$/MWh			52.1 (49.5)

el., electricity; temp., temperature

2006) as local NLP solver during pre-processing, and no local solver (pure function evaluation of the lower-bounding solution point) during the B&B. For range reduction, we use optimization-based bound tightening with trivial filtering (Gleixner et al. 2017), duality-based bound-tightening (Ryoo and Sahinidis 1995), and basic constraint propagation. The feasibility-tolerances are set to 10^{-6} and the relative optimality tolerance to 10^{-2}. All calculations are conducted on an Intel® Core™ i5-3570 CPU with 3.4 GHz running Fedora Linux 30.

Table 4 Problem statistics and solution times with a time limit of 24 h when using McCormick relaxations or the relaxations presented herein for the IAPWS-IF97 functions, as well as when using the very simple ideal model of Bongartz and Mitsos (2017)

Problem	Func.	IAPWS-IF97 model				Ideal model	
		McCormick		Proposed		McCormick	
		CPU	Iter.	CPU	Iter.	CPU	Iter.
CS1, $\dot{W}_{\text{net}}$	10	0.60 s	557	0.08 s	17	0.03 s	3
CS1, LCOE	10	4.36 s	2845	0.32 s	195	0.05 s	19
CS2, $\dot{W}_{\text{net}}$	19	>24 h (gap 4×10^{11})		0.939 s	499	0.10 s	129
CS2, LCOE	19	>24 h (gap 92%)		14.5 s	7353	0.78 s	1109
CS3, $\dot{W}_{\text{net}}$	31	>24 h (gap undef.)		9.90 min	1.6×10^5	2.92 s	2199
CS3, LCOE	31	>24 h (gap 92%)		>24 h (gap 4%)		>24 h (gap 2%)	

CS, Case Study; CPU, CPU time; Func., number of occurrences of IAPWS-IF97 functions; Iter., number of B&B iterations; McCormick, McCormick relaxations; Proposed, relaxations developed herein; gap, remaining relative optimality gap at CPU time limit; undef., gap not defined because the lower bound remained at minus infinity

4.2 Numerical results

The optimal solution points and objective values of the three case studies are summarized in Tables 1, 2 and 3. To avoid confusion with the symbols introduced in the previous sections, the stream indices are prefixed by S in the tables. For comparison, the results with the very simple ideal thermodynamic models of Bongartz and Mitsos (2017) are also shown. Note that these deviate slightly from the original values (Bongartz and Mitsos 2017) when minimizing LCOE because the cooling water temperature in the condenser needed to be modified for the chosen condenser pressure to remain feasible when using the IAPWS-IF97. While the optimal objective values differ by less than 6% between the results with the IAPWS-IF97 model and the results with the ideal model, the optimal solution points and hence the optimal cycle designs differ significantly. For example, for Case Study 1, at the solution for maximum power output the cycle pressure p_{S2} is at its upper bound (note that the bound $p_{S2} \leq 10\,\text{MPa}$ was chosen for consistency with the original problem of Bongartz and Mitsos (2017); therein, it was chosen because of the simplistic model that was expected to perform best at moderate pressures) when using the IAPWS-IF97 model, whereas with the simple model, it is approximately 50% lower (cf. Table 1). This highlights the importance of using detailed thermodynamic models for the design of steam power cycles.

When using McCormick relaxations for the IAPWS-IF97 functions, only Case Study 1 can be solved (rather quickly) within the given time limit of 24 h (see Table 4). For Case Studies 2 and 3, the global solution and thus the correct final upper bound (UBD; all problems are cast as minimization problems when implementing them for MAiNGO) is found during pre-processing as well, but the lower bound (LBD) barely improves from the values for the root node, resulting in large or even undefined (for $\dot{W}_{\text{net}}$ in Case Study 3; the lower bound remains at minus

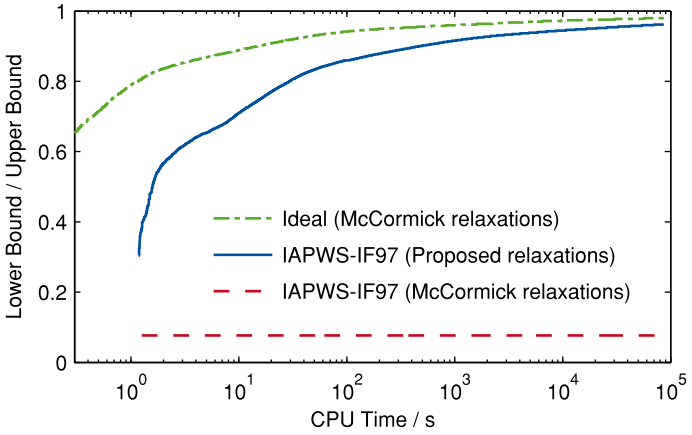

Fig. 6 When minimizing the levelized cost of electricity in Case Study 3 (dual-pressure cycle), MAiNGO closes the optimality gap between the lower and upper objective bounds significantly faster with the proposed relaxations of IAPWS-IF97 functions than with McCormick relaxations and gets closer to the performance with ideal thermodynamics (Bongartz and Mitsos 2017)

infinity, i.e., no valid overall lower bound was identified) relative optimality gaps, defined as (UBD − LBD)/UBD. Additionally, numerous numerical difficulties are encountered by CLP when solving the linear lower bounding problems (e.g., false infeasibility claims or solution points that are not actually feasible). These are likely due to the extremely large function values encountered outside the physical domains as well as the very weak relaxations (cf. Sect. 3).

When using the proposed relaxations, the solution of Case Study 1 takes almost 97% less B&B iterations and 86% (for $\dot{W}_{\text{net}}$) to 95% (for LCOE) less CPU time than the solution with McCormick relaxations. Unlike with McCormick relaxations, Case Study 2 can be solved quickly as well with either objective. For Case Study 3, only the problem of maximizing $\dot{W}_{\text{net}}$ can be solved to the desired relative optimality tolerance of 1% within a few minutes, while the minimization of LCOE terminates at the CPU time limit of 24 h with a remaining relative optimality gap of 3.7%. Nevertheless, the optimality gap is closed much faster with the proposed relaxations than with McCormick relaxations (cf. Fig. 6). Furthermore, this problem can not be solved to the desired accuracy with the current version of MAiNGO when using the very simple ideal model of Bongartz and Mitsos (2017) either (cf. Table 4 and Fig. 6). This indicates that the difficulty with this problem is not purely due to the complexity or relaxations of the IAPWS-IF97.

5 Conclusion

We have derived relaxations for relevant functions from the IAPWS-IF97 that are orders of magnitude tighter than those obtained from general purpose methods like the McCormick technique. To derive the relaxations, the functions were modified

outside their original domains in regions where the model has no physical meaning but where evaluation is required during global optimization. The functions were then analyzed for monotonicity properties to construct tight range bounds and (componentwise) convexity properties to construct tight convex and concave relaxations using variants of the αBB method as well as methods for relaxation of componentwise convex or concave functions.

The relaxations were tested on three bottoming cycles for combined cycle power plants of increasing complexity. For all but the simplest example, global optimization of the cycle design for either power output or levelized cost of electricity was not possible within reasonable computational time with McCormick relaxations but only with the relaxations developed herein. For the largest cycle, the minimization of the levelized cost of electricity could not be solved to the desired accuracy with the proposed relaxations either, although the optimality gap was closed much faster than with McCormick relaxations.

Future work could aim at improving the relaxations even further, which is in principle possible because the proposed relaxations are no envelopes, except for some of the univariate functions. Beyond better relaxations of the functions considered herein, compositions could be considered (e.g., $h_1(p,s) := h_1(p, T_1(p,s))$), given the fact that good relaxations for intrinsic functions do not always lead to good relaxations of the composite function (Najman and Mitsos 2019). Hence, tighter relaxations could be achieved by considering these composite functions as intrinsic functions themselves.

In addition to having tight relaxations for the thermodynamic models, computational advantages can also result from suitable modeling of the process flowsheets, especially in the context of reduced-space optimization formulations as considered herein (Bongartz and Mitsos 2017) that aim at a sequential evaluation of large parts of the process model. To this end, it would be beneficial to enable the use of the backward equations $T_2(p,h)$ and $T_2(p,s)$ from the IAPWS-IF97. These would allow for more freedom in the way the flowsheet is modeled, and for example allow to eliminate more optimization variables and equality constraints from the optimization problem and move them into the flowsheet evaluation, thus leading to a smaller problem and potentially further reduced runtime. When developing relaxations for $T_2(p,h)$ and $T_2(p,s)$, care needs to be taken to handle the discontinuities induced by the piecewise definition of the functions. Similar difficulties arise when extending the present approach to Region 3 of the IAPWS-IF97 (Wagner et al. 2000; IAPWS 2007b), which would enable optimization of transcritical and supercritical cycles.

Finally, from a modeling perspective, it would be desirable to avoid having to specify which point in the flowsheet lies in which region of the IAPWS-IF97. This could be achieved either by using integer variables to let the optimizer choose between subregions, or by considering functions with piecewise definition over the regions, e.g., a function $T(p, h)$ that consists of the respective functions in the different regions. Similar approaches have already been used for local dynamic optimization (Tică et al. 2012; Åberg et al. 2017), but in conjunction with simplifications (and smoothing) of the IAPWS-IF97. In analogy to the present approach, the functions could also be kept unchanged where they have physical meaning and instead be analyzed to derive tighter relaxations.

Acknowledgements Open Access funding provided by Projekt DEAL. We would like to thank Benoît Chachuat (Imperial College London) for providing MC++ and supporting us in extending it. Furthermore, we thank David Zanger (AVT.SVT, RWTH Aachen University) for implementing the functions of the original IAPWS-IF97 model. This work has received funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) "Improved McCormick Relaxations for the efficient Global Optimization in the Space of Degrees of Freedom" MI 1851/4-1. We gratefully acknowledge additional funding by the German Federal Ministry of Education and Research (BMBF) within the "Kopernikus Project P2X: Flexible use of renewable resources – exploration, validation and implementation of 'Power-to-X' concepts".

Appendix 1: Physical domains and box domains

As an example for a straightforward case, consider the function $h_1(p,T)$. Wagner et al. (2000) specify its domain as $\mathcal{P}_{h_1(p,T)} = \{(p,T)|273.15\,\mathrm{K} \le T \le 623.15\,\mathrm{K}, p_s(T) \le p \le 100\,\mathrm{MPa}\}$ (cf. also Fig. 1). We obtain the box domain $\mathcal{B}_{h_1(p,T)} = [p_1^{\min}, p_1^{\max}] \times [T_1^{\min}, T_1^{\max}] = [\min_{T\in[T_1^{\min},T_1^{\max}]} p_s(T), 100\,\mathrm{MPa}] \times [273.15, 623.15]\,\mathrm{K}$, where $p_1^{\min} = \min_{T\in[T_1^{\min},T_1^{\max}]} p_s(T) = 611.2127 \times 10^{-6}\,\mathrm{MPa}$ is also given by Wagner et al. (2000). The physical domain can be expressed as $\mathcal{P}_{h_1(p,T)} = \{(p,T) \in \mathcal{B}_{h_1(p,T)} | p_s(T) \le p\} = \{(p,T) \in \mathcal{B}_{h_1(p,T)} | T_s(p) \ge T\}$.

As an example for a more involved case, consider the function $T_1(p,h)$ which is part of the so-called *backward equations* (Wagner et al. 2000). It is related to the function $h_1(p,T)$ in the sense that for any fixed $\hat{p} \in [p_1^{\min}, p_1^{\max}]$, the function $\hat{T}_1(h) := T_1(\hat{p},h)$ is intended to be the inverse of the function $\hat{h}_1(T) := h_1(\hat{p},T)$. Although this inversion was not done analytically but $T_1(p,h)$ was fitted to data separately from $h_1(p,T)$, we can use this relation to obtain the physical domain of $T_1(p,h)$ implied by Wagner et al. (2000) as $\mathcal{P}_{T_1(p,h)} = \{(p,h)|p \in [p_1^{\min}, p_1^{\max}], h_1^{\mathrm{lower}}(p) \le h \le h_1^{\mathrm{upper}}(p)\}$, where

$$h_1^{\mathrm{lower}}(p) := \min_T \; h_1(p,T) \quad \text{s.t. } (p,T) \in \mathcal{P}_{h_1(p,T)}, \tag{22}$$

$$h_1^{\mathrm{upper}}(p) := \max_T \; h_1(p,T) \quad \text{s.t. } (p,T) \in \mathcal{P}_{h_1(p,T)}. \tag{23}$$

Table 5 Bounds of the box domains of the IAPWS-IF97 functions considered in this study. For a function $f_i(x,z)$, the box domain is $\mathcal{B}_{f_i(x,z)} := [x_i^{\min}, x_i^{\max}] \times [z_i^{\min}, z_i^{\max}]$

Bound	Unit	Value	Source
$p_1^{\min}$	MPa	611.2127×10^{-6}	Def. of Region 1
$p_1^{\max}$	MPa	100	Def. of Region 1
$T_1^{\min}$	K	273.15	Def. of Region 1
$T_1^{\max}$	K	623.15	Def. of Region 1
$h_1^{\min}$	kJ/kg	−0.04158783	Min. of $h_1(p,T)$ over $\mathcal{P}_{h_1(p,T)}$
$h_1^{\max}$	kJ/kg	1671.023	Max. of $h_1(p,T)$ over $\mathcal{P}_{h_1(p,T)}$
$s_1^{\min}$	kJ/(kg K)	−0.008582287	Min. of $s_1(p,T)$ over $\mathcal{P}_{s_1(p,T)}$
$s_1^{\max}$	kJ/(kg K)	3.778281	Max. of $s_1(p,T)$ over $\mathcal{P}_{s_1(p,T)}$
$p_2^{\min}$	MPa	611.2127×10^{-6}	Def. herein (cf. Sect. 2.1)
$p_2^{\max}$	MPa	100	Def. of Region 2
$T_2^{\min}$	K	273.15	Def. of Region 2
$T_2^{\max}$	K	1073.15	Def. of Region 2
$h_2^{\min}$	kJ/kg	2500.82500	Min. of $h_2(p,T)$ over $\mathcal{P}_{h_2(p,T)}$
$h_2^{\max}$	kJ/kg	4160.66300	Max. of $h_2(p,T)$ over $\mathcal{P}_{h_2(p,T)}$
$s_2^{\min}$	kJ/(kg K)	5.048097	Min. of $s_2(p,T)$ over $\mathcal{P}_{s_2(p,T)}$
$s_2^{\max}$	kJ/(kg K)	11.92106	Max. of $s_2(p,T)$ over $\mathcal{P}_{s_2(p,T)}$
$p_3^{\min}$	MPa	623.15	Def. of Region 3
$T_3^{\min}$	K	16.5292	Def. of Region 3
$p_{4-1/2}^{\min}$	MPa	611.2127×10^{-6}	Def. of Region 4-1/2 (cf. Sect. 3)
$p_{4-1/2}^{\max}$	MPa	16.5292	Def. of Region 4-1/2 (cf. Sect. 3)
$T_{4-1/2}^{\min}$	K	273.15	Def. of Region 4-1/2 (cf. Sect. 3)
$T_{4-1/2}^{\max}$	K	623.15	Def. of Region 4-1/2 (cf. Sect. 3)
$y_{4-1/2}^{\min}$	–	0	Def. of Region 4-1/2 (cf. Sect. 3)
$y_{4-1/2}^{\min}$	–	1	Def. of Region 4-1/2 (cf. Sect. 3)
$h_{4-1/2}^{\max}$	kJ/kg	−0.04158783	Min. of $h_{4-1/2}(p,y)$ over $\mathcal{P}_{h_{4-1/2}(p,y)}$
$h_{4-1/2}^{\max}$	kJ/kg	2803.285	Max. of $h_{4-1/2}(p,y)$ over $\mathcal{P}_{h_{4-1/2}(p,y)}$
$s_{4-1/2}^{\min}$	kJ/(kg K)	−0.0001545496	Min. of $s_{4-1/2}(p,y)$ over $\mathcal{P}_{s_{4-1/2}(p,y)}$
$s_{4-1/2}^{\max}$	kJ/(kg K)	9.155759	Max. of $s_{4-1/2}(p,y)$ over $\mathcal{P}_{s_{4-1/2}(p,y)}$

Def., Definition; Min., Global minimization, Max., Global maximization

The parametric optimization problems in (22) and (23) can be solved analytically since $h_1(p,T)$ is monotonically increasing with respect to T (cf. Sect. 3.3) and we obtain

$$h_1^{\text{lower}}(p) = h_1(p, T_1^{\min})$$
$$h_1^{\text{upper}}(p) = \begin{cases} h_1(p, T_s(p)), & \text{if } p \le p_3^{\min} \\ h_1(p, T_1^{\max}), & \text{otherwise.} \end{cases}$$

Table 6 Physical domains of the IAPWS-IF97 functions considered in this study. For a function $f_i(x,z)$, $\mathcal{B}_{f_i(x,z)}$ denotes its box domain according to Table 5

Function	Physical domain
$h_1(p,T)$	$\{(p,T) \in \mathcal{B}_{h_1(p,T)} \mid p \geq p_s(T)\}$
$s_1(p,T)$	$\{(p,T) \in \mathcal{B}_{h_1(p,T)} \mid p \geq p_s(T)\}$
$T_1(p,h)$	$\{(p,h) \in \mathcal{B}_{T_1(p,h)} \mid h_1^{\text{lower}}(p) \leq h \leq h_1^{\text{upper}}(p)\}$, $h_1^{\text{lower}}(p) := h_1(p, T_1^{\min})$, $h_1^{\text{upper}}(p) := \begin{cases} h_1(p, T_s(p)), & \text{if } p \leq p_3^{\min} \\ h_1(p, T_1^{\max}), & \text{if } p > p_3^{\min} \end{cases}$
$T_1(p,s)$	$\{(p,s) \in \mathcal{B}_{T_1(p,s)} \mid s_1^{\text{lower}}(p) \leq s \leq s_1^{\text{upper}}(p)\}$, $s_1^{\text{lower}}(p) := s_1(p, T_1^{\min})$, $s_1^{\text{upper}}(p) := \begin{cases} s_1(p, T_s(p)), & \text{if } p \leq p_3^{\min} \\ s_1(p, T_1^{\max}), & \text{if } p > p_3^{\min} \end{cases}$
$h_2(p,T)$	$\{(p,T) \in \mathcal{B}_{h_2(p,T)} \mid p \leq p_2^{\text{upper}}(T)\}$, $p_2^{\text{upper}}(T) := \begin{cases} p_s(T), & \text{if } T \leq T_3^{\min} \\ p_{\text{B23}}(T), & \text{if } T > T_3^{\min} \end{cases}$
$s_2(p,T)$	$\{(p,T) \in \mathcal{B}_{s_2(p,T)} \mid p \leq p_2^{\text{upper}}(T)\}$
$h_{4-1/2}(p,y)$	$\{(p,y) \in \mathcal{B}_{h_{4-1/2}(p,y)}\}$
$s_{4-1/2}(p,y)$	$\{(p,y) \in \mathcal{B}_{s_{4-1/2}(p,y)}\}$
$y_{4-1/2}(p,h)$	$\{(p,h) \in \mathcal{B}_{y_{4-1/2}(p,h)} \mid h_{4-1/2}^{\text{liq}}(p) \leq h \leq h_{4-1/2}^{\text{vap}}(p)\}$
$y_{4-1/2}(p,s)$	$\{(p,s) \in \mathcal{B}_{y_{4-1/2}(p,s)} \mid s_{4-1/2}^{\text{liq}}(p) \leq s \leq s_{4-1/2}^{\text{vap}}(p)\}$

Finally, for the box domain $\mathcal{B}_{T_1(p,h)} = [p_1^{\min}, p_1^{\max}] \times [h_1^{\min}, h_1^{\max}]$ we obtain $h_1^{\min}$ and $h_1^{\max}$ by globally minimizing $h_1^{\text{lower}}(p)$ and maximizing $h_1^{\text{upper}}(p)$, respectively, over $[p_1^{\min}, p_1^{\max}]$. The physical and box domains of the remaining functions can be found in Tables 5 and 6.

Appendix 2: Modification outside the physical domain

By means of example, consider the function $h_2(p,T)$. Its physical domain $\mathcal{P}_{h_2(p,T)} = \{(p,T) \in \mathcal{B}_{h_2(p,T)} \mid p \leq p_2^{\text{upper}}(T)\}$, where

$$p_2^{\text{upper}}(T) = \begin{cases} p_s(T), & \text{if } T \leq T_3^{\min} \\ p_{\text{B23}}(T), & \text{otherwise,} \end{cases} \tag{24}$$

is shown in Fig. 2a, where the solid line denotes the points $(p_2^{\text{upper}}(\tilde{T}), \tilde{T})$. When evaluating $h_2(p,T)$ on the entire box domain $\mathcal{B}_{h_2(p,T)}$, we observe undesired peaks of very large magnitude for points $(\tilde{p}, \tilde{T})$ with $\tilde{p} > p_2^{\text{upper}}(\tilde{T})$ (cf. Fig. 2b). These peaks destroy the monotonicity properties that $h_2(p,T)$ exhibits on its physical domain

Table 7 Intermediate modifications of IAPWS-IF97 functions outside their physical domains to enable the construction of tight relaxations. $k_i, i = 1, \ldots, 12$ are parameters

Function	Intermediate modified functions
$h_1(p,T)$	$h_1^{\text{int}}(p,T) := \begin{cases} h_1(p,T), & \text{if } p \geq p_s(T) \\ h_1(p_s(T),T) + \left(\frac{\partial h_1}{\partial p}\right)\Big\|_{(p_s(T),T)} \big(p - p_s(T)\big), & \text{if } p < p_s(T) \end{cases}$
$s_1(p,T)$	$s_1^{\text{int}}(p,T) := \begin{cases} s_1(p,T), & \text{if } p \geq p_s(T) \\ s_1(p_s(T),T) + \left(\frac{\partial s_1}{\partial p}\right)\Big\|_{(p_s(T),T)} \big(p - p_s(T)\big), & \text{if } p < p_s(T) \end{cases}$
$T_1(p,h)$	$T_1^{\text{int}}(p,h) := \begin{cases} T_1(p,h), & \text{if } p \geq p_3^{\min} \vee h \leq h_1(p,T_s(p)) \\ T_1(p,h_1(p,T_s(p))) + k_1\big(h - h_1(p,T_s(p))\big), & \text{otherwise} \end{cases}$
$h_2(p,T)$	$h_2^{\text{int}}(p,T) := \begin{cases} h_2(p,T), & \text{if } p \leq p_2^{\text{lim}}(T) \\ h_2(p_2^{\text{lim}}(T),T) - \left(k_5 + \frac{k_6 T}{\sqrt{p_2^{\text{lim}}(T)}}\right)\big(p - p_2^{\text{lim}}(T)\big), & \text{otherwise} \end{cases}$
	$p_2^{\text{lim}}(T) := \begin{cases} p_s(T), & \text{if } T \leq 350\,\text{K} \\ k_1 + k_2 T + k_3 T^2 + k_4 T^3, & \text{otherwise} \end{cases}$
$s_2(p,T)$	$s_2^{\text{int}}(p,T) := \begin{cases} s_2(p,T), & \text{if } T \geq T_2^{\text{lim}}(p) \\ s_2(p,T_2^{\text{lim}}(p)) + k_8\big(T - T_2^{\text{lim}}(p)\big), & \text{otherwise.} \end{cases}$
	$T_2^{\text{lim}}(p) := \begin{cases} T_s(p), & \text{if } p \leq p_3^{\min} \\ k_9 + k_{10} T + k_{11} T^2 + k_{12} T^3, & \text{otherwise} \end{cases}$

$\mathcal{P}_{h_2(p,T)}$ and would additionally lead to rather weak relaxations over boxes including these peaks.

To enable tighter relaxations, we replace the function with a suitable extrapolation in this region. However, since $p_2^{\text{upper}}(T)$ is nonsmooth (cf. (24) and Fig. 2a), we consider a relaxed physical domain $\mathcal{P}^{\text{rel}}_{h_2(p,T)} := \{(p,T) \in \mathcal{B}_{h_2(p,T)} | p \leq p_2^{\text{lim}}(T)\}$, with

$$p_2^{\text{lim}}(T) := \begin{cases} p_s(T), & \text{if } T \leq 350\,\text{K} \\ k_1 + k_2 T + k_3 T^2 + k_4 T^3, & \text{otherwise,} \end{cases} \tag{25}$$

where k_1–k_4 are chosen such that $p_2^{\text{lim}}(T)$ is continuously differentiable and $p_2^{\text{lim}}(T) \geq p_2^{\text{upper}}(T)\ \forall T \in \{\hat{T} \in [T_2^{\min}, T_2^{\max}] \mid p_2^{\text{upper}}(\hat{T}) \leq p_2^{\max}\}$ (cf. Fig. 2a). We then define the intermediate modified function

Table 8 Monotonicity guarantees for the (modified) functions from the IAPWS-IF97

Function	Monotonicity guarantees
$p_s(T)$	inc.
$T_s(p)$	inc.
$h^{\text{liq}}_{4\text{-}1/2}(p)$	inc.
$h^{\text{vap}}_{4\text{-}1/2}(p)$	inc. if $p \leq 3.0783756970\,\text{MPa}$
	dec. if $p \geq 3.0783756971\,\text{MPa}$
$s^{\text{liq}}_{4\text{-}1/2}(p)$	inc.
$s^{\text{vap}}_{4\text{-}1/2}(p)$	dec.
$p_{\text{B23}}(T)$	inc.
$T_{\text{B23}}(p)$	inc.
$h^{\text{mod}}_1(p,T)$	inc. w.r.t. T
	inc. w.r.t. p if $T \leq 510\,\text{K}$
	dec. w.r.t. p if $T \geq 614\,\text{K}$
$s^{\text{mod}}_1(p,T)$	inc. w.r.t. p
	dec. w.r.t. T if $p \geq 19\,\text{MPa} \vee T \geq 278\,\text{K}$
$T^{\text{mod}}_1(p,h)$	inc. w.r.t. h
	dec. w.r.t. p if $h \leq 1073\,\text{kJ/kg} \wedge h \leq h^{\text{liq}}_{4\text{-}1/2}(p)$
	inc. w.r.t. p if $p \leq p^{\min}_3 \wedge h \geq h^{\text{liq}}_{4\text{-}1/2}(p)$
$T^{\text{mod}}_1(p,s)$	not analyzed
$h^{\text{mod}}_2(p,T)$	inc. w.r.t. p
	dec. w.r.t. T
$s^{\text{mod}}_2(p,T)$	inc. w.r.t. p
	dec. w.r.t. T
$h^{\text{mod}}_{4\text{-}1/2}(p,y)$	inc. w.r.t. y
$s^{\text{mod}}_{4\text{-}1/2}(p,y)$	inc. w.r.t. y
$y^{\text{mod}}_{4\text{-}1/2}(p,h)$	inc. w.r.t. h
	dec. w.r.t. p if $p \leq 3\,\text{MPa} \vee h \leq 2158\,\text{kJ/kg}$
$y^{\text{mod}}_{4\text{-}1/2}(p,s)$	inc. w.r.t. s

inc., increasing; dec., decreasing

$$h_2^{\text{int}}(p,T) := \begin{cases} h_2(p,T), & \text{if } p \leq p_2^{\lim}(T) \\ h_2(p_2^{\lim}(T),T) - \left(k_5 + \frac{k_6 T}{\sqrt{p_2^{\lim}(T)}}\right)\left(p - p_2^{\lim}(T)\right), & \text{otherwise,} \end{cases} \tag{26}$$

with parameters k_5 and k_6. The extension for $p > p_2^{\lim}(T)$ is chosen such that $h_2^{\text{int}}(p,T)$ is continuous, it is increasing with respect to p and decreasing with respect to T, it is componentwise concave with respect to p, componentwise concave with respect to T

Table 9 Convexity guarantees for the (intermediate) functions from the IAPWS-IF97 that are exploited in this work

Function	Convexity guarantees
$p_s(T)$	conv.
$T_s(p)$	conc.
$h^{\text{liq}}_{4\text{-}1/2}(p)$	conc. if $p \leq 14.48\,\text{MPa}$
$h^{\text{vap}}_{4\text{-}1/2}(p)$	conc.
$s^{\text{liq}}_{4\text{-}1/2}(p)$	conc. if $p \leq 15.26\,\text{MPa}$
$s^{\text{vap}}_{4\text{-}1/2}(p)$	conv. if $p \leq 12.23\,\text{MPa}$
$p_{\text{B23}}(T)$	conv.
$T_{\text{B23}}(p)$	conc.
$h^{\text{int}}_1(p,T)$	comp. conc. w.r.t. T if $T \geq 314\,\text{K} \vee p \geq 26\,\text{MPa}$
	comp. conv. w.r.t. p if $T \geq 370\,\text{K}$
$s^{\text{int}}_1(p,T)$	comp. conv. w.r.t. p if $T \geq 319\,\text{K}$
$T^{\text{int}}_1(p,h)$	comp. conc. w.r.t. p
	comp. conc. w.r.t. h if $p \geq 16.4\,\text{MPa} \vee h \geq 166\,\text{kJ/kg}$
$T^{\text{int}}_1(p,s)$	not analyzed
$h^{\text{int}}_2(p,T)$	comp. conc. w.r.t. p if $p \leq 28.68\,\text{MPa}$
$s^{\text{int}}_2(p,T)$	comp. conv. w.r.t. p if $T \geq 794\,\text{K}$
	comp. conc. w.r.t. T
$h^{\text{int}}_{4\text{-}1/2}(p,y)$	not analyzed
$s^{\text{int}}_{4\text{-}1/2}(p,y)$	not analyzed
$y^{\text{int}}_{4\text{-}1/2}(p,h)$	not analyzed
$y^{\text{int}}_{4\text{-}1/2}(p,s)$	not analyzed

comp., componentwise; conv., convex; conc., concave

on most of the box domain, and has $\frac{\partial^2 h_2^{\text{int}}}{\partial p \partial T} \geq 0$. Furthermore, while $h_2^{\text{int}}(p,T)$ does attain values significantly below $h_2^{\min} := \min_{(p,T)\in\mathcal{P}_{h_2(p,T)}} h_2(p,T)$ for points $(\tilde{p},\tilde{T})$ with $\tilde{p} > p_2^{\text{upper}}(\tilde{T})$, the values are much smaller in magnitude than those attained by $h_2(p,T)$ in this region (cf. Fig. 2b vs. 2c). Although $h_2^{\text{int}}(p,T)$ is nonsmooth at every point $(p_2^{\text{lim}}(\tilde{T}),\tilde{T})$, the extrapolation is constructed to have a negative solution value of the maximization problem analogous to (5). Compared with the original function $h_2(p,T)$, $h_2^{\text{int}}(p,T)$ thus exhibits useful properties that can be used to construct tight range bounds and relaxations.

We then define the final modified function

$$h_2^{\text{mod}}(p,T) := \max\left(h_2^{\text{int}}(p,T), h_2^{\min}\right) \tag{27}$$

that cuts off $h_2^{\text{int}}(p,T)$ at the minimum value of $h_2(p,T)$ over $\mathcal{P}_{h_2(p,T)}$. Note that cutting off at $h_2^{\max} := \max_{(p,T)\in\mathcal{P}_{h_2(p,T)}} h_2(p,T)$ is not required because the maximum of $h_2^{\text{int}}(p,T)$ is attained in the physical domain. The graph of $h_2^{\text{mod}}(p,T)$ is shown in Fig. 2d, where the extrapolation according to (26) is shown in green and the changes induced by the max operator in (27) are shown in orange. Relaxations of $h_2^{\text{mod}}(p,T)$ can be obtained by correcting the relaxations of $h_2^{\text{int}}(p,T)$ using the rules for

relaxation of the max function (cf. Sect. 3.4). Given the smaller range of $h_2^{\mathrm{mod}}(p,T)$ compared with $h_2(p,T)$, even the convex envelope of $h_2^{\mathrm{mod}}(p,T)$ over $\mathcal{B}_{h_2(p,T)}$ or large subsets thereof would be tighter over $\mathcal{P}_{h_2(p,T)}$ than that of $h_2(p,T)$. Table 7 summarized the intermediate modifications of the remaining bivariate functions.

Appendix 3: Monotonicity and range bounds

As a simple example, consider the univariate function $p_\mathrm{s}(T)$. The function is monotonically increasing on its entire domain (cf. Table 8), such that for any $[T^\mathrm{L},T^\mathrm{U}] \subseteq [T_4^\mathrm{min},T_4^\mathrm{max}]$ we obtain exact range bounds as $p_\mathrm{s}\left([T^\mathrm{L},T^\mathrm{U}]\right) = \left[p_\mathrm{s}(T^\mathrm{L}),p_\mathrm{s}(T^\mathrm{U})\right]$.

As an example for a more involved case, we consider the function $h_1^{\mathrm{mod}}(p,T)$. For $\mathcal{P} := [p^\mathrm{L},p^\mathrm{U}] \subseteq [p_1^\mathrm{min},p_1^\mathrm{max}]$, $\mathcal{T} := [T^\mathrm{L},T^\mathrm{U}] \subseteq [T_1^\mathrm{min},T_1^\mathrm{max}]$, we obtain $h_1^{\mathrm{mod}}(\mathcal{P}\times\mathcal{T}) \subseteq \left[\hat{h}^\mathrm{L},\hat{h}^\mathrm{U}\right]$ with

$$\hat{h}^\mathrm{U} = \begin{cases} h_1^{\mathrm{mod}}(p^\mathrm{U},T^\mathrm{U}), & \text{if } T^\mathrm{U} \le 510\,\mathrm{K} \\ h_1^{\mathrm{mod}}(p^\mathrm{L},T^\mathrm{U}), & \text{if } T^\mathrm{U} \ge 614\,\mathrm{K} \\ \max\left(h_1^{\mathrm{mod}}(p^\mathrm{L},T^\mathrm{U}),h_1^{\mathrm{mod}}(p^\mathrm{U},T^\mathrm{U})\right), & \text{if } T^\mathrm{U} \in (510,614)\,\mathrm{K}, \end{cases}$$

$$\hat{h}^\mathrm{L} = \begin{cases} h_1^{\mathrm{mod}}(p^\mathrm{L},T^\mathrm{L}), & \text{if } T^\mathrm{L} \le 510\,\mathrm{K} \\ h_1^{\mathrm{mod}}(p^\mathrm{U},T^\mathrm{L}), & \text{if } T^\mathrm{L} \ge 614\,\mathrm{K} \\ h_1^{\mathrm{mod}}(p^\mathrm{L},T^\mathrm{L}), & \text{if } T^\mathrm{L} \in (510,614)\,\mathrm{K} \wedge \left.\frac{\partial h_1}{\partial p}\right|_{(p^\mathrm{L},T^\mathrm{L})} \ge 0 \\ h_1^{\mathrm{mod}}(p^\mathrm{U},T^\mathrm{L}), & \text{if } T^\mathrm{L} \in (510,614)\,\mathrm{K} \wedge \left.\frac{\partial h_1}{\partial p}\right|_{(p^\mathrm{U},T^\mathrm{L})} \le 0 \\ \max\left(\mathrm{IE}^\mathrm{L}_{h_1(p,T)}\left(\widehat{\mathcal{PT}}\right),\right. & \\ \left. h_1^{\mathrm{mod}}(p^\mathrm{L},510\,\mathrm{K})\right), & \text{otherwise,} \end{cases}$$

where $\mathrm{IE}^\mathrm{L}_{h_1(p,T)}\left(\widehat{\mathcal{PT}}\right)$ denotes a lower bound for $h_1(p,T)$ over the set $\widehat{\mathcal{PT}} := [\max(p^\mathrm{L},p_\mathrm{s}(T^\mathrm{L})),p^\mathrm{U}] \times \left\{T^\mathrm{L}\right\}$ computed via natural interval extensions. Since $h_1^{\mathrm{mod}}(p,T)$ is increasing with respect to T (cf. Table 8), the maximum and minimum over $\mathcal{P}\times\mathcal{T}$ are attained at T^U and T^L, respectively. For $T^\mathrm{U} \le 510\,\mathrm{K}$ or $T^\mathrm{U} \ge 614\,\mathrm{K}$, $h_1^{\mathrm{mod}}(p,T^\mathrm{U})$ is monotonic in p as well and the maximum is thus at p^U or p^L, respectively. For $T^\mathrm{U} \in (510,614)\,\mathrm{K}$, $h_1^{\mathrm{mod}}(p,T^\mathrm{U})$ is not monotonic in p, but since it is componentwise convex with respect to p in this region (cf. Table 9), the maximum is attained at either p^L or p^U. Similarly, for $T^\mathrm{L} \le 510\,\mathrm{K}$ or $T^\mathrm{L} \ge 614\,\mathrm{K}$, $h_1^{\mathrm{mod}}(p,T^\mathrm{L})$ is monotonic in p as well and the minimum is thus attained at p^U or p^L, respectively. However, for $T^\mathrm{L} \in (510,614)\,\mathrm{K}$, $h_1^{\mathrm{mod}}(p,T^\mathrm{L})$ is not monotonic in p, and because it is componentwise convex, the minimum could be attained at any $p \in \mathcal{P}$. In this case, we can exploit componentwise convexity with respect to p to conclude that if the partial derivative is non-negative at p^L or non-positive at p^U, the minimum must lie at that value of p. If this is not the case, the minimum can lie anywhere in $\widehat{\mathcal{PT}} = [\max(p^\mathrm{L},p_\mathrm{s}(T^\mathrm{L})),p^\mathrm{U}] \times \left\{T^\mathrm{L}\right\}$ (note that it cannot lie below $p_\mathrm{s}(T^\mathrm{L})$ because the function is linear with respect to p for $p < p_\mathrm{s}(T)$ (cf. Table 7) and we are

in the case where $\left.\frac{\partial h_1}{\partial p}\right|_{(p^{\mathrm{L}},T^{\mathrm{L}})} < 0$). Since $\widehat{\mathcal{PT}} \subset \mathcal{P}_{h_1(p,T)}$, we have $h_1^{\mathrm{mod}}(\tilde{p},\tilde{T}) = h_1(\tilde{p},\tilde{T})$ for every $(\tilde{p},\tilde{T}) \in \widehat{\mathcal{PT}}$ and thus $h_1^{\mathrm{mod}}(\tilde{p},\tilde{T})$ has a factorable representation over $\widehat{\mathcal{PT}}$ and we can use natural interval extensions from FILIB++ to obtain an underestimation of the minimum function value. Another, potentially tighter, lower bound can be obtained by exploiting the monotonicity with respect to T that implies that a lower bound over $\mathcal{P} \times \{\hat{T}\}$ is a valid lower bound over $\mathcal{P} \times \{T^{\mathrm{L}}\}$ for every $\hat{T} \leq T^{\mathrm{L}}$. In particular, this holds for $\hat{T} = 510\,\mathrm{K}$, for which we know that the lower bound is attained at p^{L}.

Appendix 4: Convexity and relaxations

As an example for a univariate function, we consider the function $h_{4\text{-}1/2}^{\mathrm{liq}}(p)$, which is defined on $\mathcal{P}_{h_{4\text{-}1/2}^{\mathrm{liq}}(p)} = [611.2127 \times 10^{-6}, 16.5292]\,\mathrm{MPa}$ but is convex only on $[611.2127 \times 10^{-6}, 14.48]\,\mathrm{MPa}$ (cf. Table 9). By globally maximizing the second derivative of $h_{4\text{-}1/2}^{\mathrm{liq}}(p)$, we obtain $\alpha := 0.5 \times 1.0592301\,\mathrm{kJ/(kgMPa^2)} \geq 0.5 \times \max_{p \in \mathcal{P}_{h_{4\text{-}1/2}^{\mathrm{liq}}(p)}} \frac{d^2 h_{4\text{-}1/2}^{\mathrm{liq}}}{dp^2}$. Given a non-degenerate interval $[p^{\mathrm{L}}, p^{\mathrm{U}}]$ and range bounds $[h^{\mathrm{L}}, h^{\mathrm{U}}]$, we construct a convex relaxation as the secant of a concave underestimator based on (12) as

$$h_{4\text{-}1/2}^{\mathrm{liq,cv,u}}(p) := \begin{cases} h_{4\text{-}1/2}^{\mathrm{liq}}(p^{\mathrm{L}}) \\ \quad + \dfrac{h_{4\text{-}1/2}^{\mathrm{liq}}(p^{\mathrm{U}}) - h_{4\text{-}1/2}^{\mathrm{liq}}(p^{\mathrm{L}})}{p^{\mathrm{U}} - p^{\mathrm{L}}}\left(p - p^{\mathrm{L}}\right), & \text{if } p^{\mathrm{U}} \leq 14.48\,\mathrm{MPa} \\ \max\left(h^{\mathrm{L}},\ h_{4\text{-}1/2}^{\mathrm{liq}}(p^{\mathrm{L}}) - \alpha\left(\dfrac{p^{\mathrm{U}} - p^{\mathrm{L}}}{2}\right)^2\right. & \\ \quad \left. + \dfrac{h_{4\text{-}1/2}^{\mathrm{liq}}(p^{\mathrm{U}}) - h_{4\text{-}1/2}^{\mathrm{liq}}(p^{\mathrm{L}})}{p^{\mathrm{U}} - p^{\mathrm{L}}}\left(p - p^{\mathrm{L}}\right)\right), & \text{otherwise,} \end{cases} \tag{28}$$

and a concave relaxation based on (13) as

$$h_{4\text{-}1/2}^{\mathrm{liq,cc,o}}(p) := \begin{cases} h_{4\text{-}1/2}^{\mathrm{liq}}(p), & \text{if } p^{\mathrm{U}} \leq 14.48\,\mathrm{MPa} \\ \min\left(h^{\mathrm{U}}, h_{4\text{-}1/2}^{\mathrm{liq}}(p) - \alpha\left(p - p^{\mathrm{L}}\right)\left(p - p^{\mathrm{U}}\right)\right), & \text{otherwise.} \end{cases} \tag{29}$$

The max and min functions in (28) and (29) potentially tighten the relaxation in case we do not have an envelope anyway. The resulting relaxations are orders of magnitude tighter than those obtained by applying standard McCormick relaxations to the factorable representation of $h_{4\text{-}1/2}^{\mathrm{liq}}(p)$ (cf. Fig. 3a, c). Furthermore, the convex relaxation (28) obtained from the αBB variant (12) by Hasan (2018) is significantly tighter than that obtained from the regular αBB version (10) (cf. Fig. 3b, c). This is due to the fact that the function itself is *almost* concave in the sense that the maximum of the second derivative is much smaller in magnitude than the minimum.

References

Åberg M, Windahl J, Runvik H, Magnusson F (2017) Optimization-friendly thermodynamic properties of water and steam. In: Proceedings of the 12th international modelica conference, Prague, Czech Republic, May 15–17, 2017. Linköping University Electronic Press, Linköpings Universitet, pp 449–458

Adjiman CS, Dallwig S, Floudas CA, Neumaier A (1998) A global optimization method, αBB, for general twice-differentiable constrained NLPs- I. Theoretical advances. Comput Chem Eng 22:1137–1158

Ahadi-Oskui T, Alperin H, Nowak I, Cziesla F, Tsatsaronis G (2006) A relaxation-based heuristic for the design of cost-effective energy conversion systems. Energy 31:1346–1357

Ahadi-Oskui T, Vigerske S, Nowak I, Tsatsaronis G (2010) Optimizing the design of complex energy conversion systems by branch and cut. Comput Chem Eng 34:1226–1236

Androulakis IP, Maranas CD, Floudas CA (1995) αBB: a global optimization method for general constrained nonconvex problems. J Glob Optim 7:337–363

Bendtsen C, Stauning O (2012) FADBAD++, a flexible C++ package for automatic differentiation. Version 2.1. http://www.fadbad.com. Accessed 18 Oct 2016

Bongartz D, Mitsos A (2017) Deterministic global optimization of process flowsheets in a reduced space using McCormick relaxations. J Glob Optim 69:761–796

Bongartz D, Najman J, Sass S, Mitsos A (2018) MAiNGO–McCormick-based Algorithm for mixed-integer Nonlinear Global Optimization. Process Systems Engineering (AVT.SVT), RWTH Aachen University. http://permalink.avt.rwth-aachen.de/?id=729717. Accessed 25 Oct 2019

Bruno J, Fernandez F, Castells F, Grossmann I (1998) A rigorous MINLP model for the optimal synthesis and operation of utility plants. Chem Eng Res Des 76:246–258

Chachuat B, Houska B, Paulen R, Perić N, Rajyaguru J, Villanueva M (2015) Set-theoretic approaches in analysis, estimation and control of nonlinear systems. IFAC-PapersOnLine 48:981–995. https://omega-icl.github.io/mcpp/. Accessed 25 Oct 2019

Falk JE, Soland RM (1969) An algorithm for separable nonconvex programming problems. Manag Sci 15:550–569

Forrest JJ, Vigerske S, Ralphs T, Hafer L, Fasano JP, Santos HG, Saltzman M, Gassmann H, Kristjansson B, King A (2019) COIN-OR linear programming solver. https://github.com/coin-or/Clp. Accessed 25 Oct 2019

Gleixner AM, Berthold T, Müller B, Weltge S (2017) Three enhancements for optimization-based bound tightening. J Glob Optim 67:731–757

Hasan FMM (2018) An edge-concave underestimator for the global optimization of twice-differentiable nonconvex problems. J Glob Optim 71:735–752

International Energy Agency (2019) Electricity information: overview. https://webstore.iea.org/electricity-information-2019. Accessed 6 Oct 2019

Khan KA, Barton PI (2015) A vector forward mode of automatic differentiation for generalized derivative evaluation. Optim Method Softw 30:1185–1212

Koch C, Cziesla F, Tsatsaronis G (2007) Optimization of combined cycle power plants using evolutionary algorithms. Chem Eng Process 46:1151–1159

Lerch M, Tischler G, Wolff von Gudenberg J, Hofschuster W, Krämer W (2011) FILIB++ Interval Library (V 3.0.2). http://www2.math.uni-wuppertal.de/wrswt/software/filib.html. Accessed 25 Oct 2019

Locatelli M, Schoen F (2013) Global optimization: theory, algorithms, and applications. MOS-SIAM, Philadelphia

Luo X, Zhang B, Chen Y, Mo S (2011) Modeling and optimization of a utility system containing multiple extractions steam turbines. Energy 36:3501–3512

Manassaldi JI, Mussati SF, Scenna NJ (2011) Optimal synthesis and design of heat recovery steam generation (HRSG) via mathematical programming. Energy 36:475–485

Manassaldi JI, Arias AM, Scenna NJ, Mussati MC, Mussati SF (2016) A discrete and continuous mathematical model for the optimal synthesis and design of dual pressure heat recovery steam generators coupled to two steam turbines. Energy 103:807–823

McCormick G (1976) Computability of global solutions to factorable nonconvex programs: Part I—Convex underestimating problems. Math Prog 10:147–175

Meyer CA, Floudas CA (2005) Convex envelopes for edge-concave functions. Math Prog 103:207–224

Mistry M, Misener R (2016) Optimising heat exchanger network synthesis using convexity properties of the logarithmic mean temperature difference. Comput Chem Eng 94:1–17
Nadir M, Ghenaiet A (2015) Thermodynamic optimization of several (heat recovery steam generator) HRSG configurations for a range of exhaust gas temperatures. Energy 86:685–695
Najman J, Mitsos A (2016) Convergence order of McCormick relaxations of LMTD function in heat exchanger networks. In: Kravanja Z, Bogataj M (eds) Proceedings of the 26th European symposium on computer aided process engineering—ESCAPE 26, pp 1605–1610
Najman J, Mitsos A (2019) On tightness and anchoring of McCormick and other relaxations. J Glob Optim 74:677–703
Najman J, Bongartz D, Mitsos A (2019a) Convex relaxations of componentwise convex functions. Comput Chem Eng 130:106527
Najman J, Bongartz D, Mitsos A (2019b) Relaxations of thermodynamic property and costing models in process engineering. Comput Chem Eng 130:106571
Nowak I, Vigerske S (2008) LaGO: a (heuristic) branch and cut algorithm for nonconvex MINLPs. Cent Eur J Oper Res 16:127–138
Podolski WF, Schmalzer DK, Conrad V, Lowenhaupt DE, Winschel RA, Klunder EB, McIlvried III HG, Ramezan M, Stiegel GJ, Srivastava RD, Winslow J, Loftus PJ, Benson CE, Wheeldon JM, Krumpelt M, Smith FL (2008) Energy resources, conversion, and utilization. In: Green DW, Perry RH (eds) Perry's chemical engineers' handbook. McGraw-Hill, New York, pp 24–1 – 24–57
Rockafellar RT (1970) Convex analysis. Princeton University Press, Princeton
Ryoo HS, Sahinidis NV (1995) Global optimization of nonconvex NLPs and MINLPs with applications in process design. Comput Chem Eng 19:551–566
Savola T, Tveit TM, Fogelholm CJ (2007) A MINLP model including the pressure levels and multiperiods for CHP process optimisation. Appl Therm Eng 27:1857–1867
Schweidtmann AM, Huster WR, Lüthje JT, Mitsos A (2019) Deterministic global process optimization: accurate (single-species) properties via artificial neural networks. Comput Chem Eng 121:67–74
Smith EM, Pantelides CC (1997) Global optimisation of nonconvex MINLPs. Comput Chem Eng 21:S791–S796
Tardella F (2004) On the existence of polyhedral convex envelopes. In: Floudas CA, Pardalos P (eds) Frontiers in global optimization. Kluwer Academic Publishers, Dordrecht, pp 563–573
Tawarmalani M, Sahinidis NV (2002) Convexifixation and global optimization in continuous and mixed-integer nonlinear programming. Kluwer Academic Publishers, Dordrecht
The International Association for the Properties of Water and Steam (2007a) IAPWS R7-97(2012)—Revised release on the IAPWS industrial formulation 1997 for the thermodynamic properties of water and steam. http://iapws.org/relguide/IF97-Rev.html. Accessed 29 Aug 2019
The International Association for the Properties of Water and Steam (2007b) IAPWS R7-97(2012)—Revised supplementary release on backward equations for the functions T(p,h), v(p,h) and T(p,s), v(p,s) for Region 3 of the IAPWS Industrial Formulation 1997 for the thermodynamic properties of water and steam. http://iapws.org/relguide/IF97-Rev.html. Accessed 26 Sept 2019
Tică A, Guéguen H, Dumur D, Faille D, Davelaar F (2012) Design of a combined cycle power plant model for optimization. Appl Energy 98:256–265
Tsoukalas A, Mitsos A (2014) Multivariate McCormick relaxations. J Glob Optim 59:633–662
Wächter A, Biegler LT (2006) On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. Math Prog 106:25–57
Wagner W, Pruss A (2002) The IAPWS formulation 1995 for the thermodynamic properties of ordinary water substance for general and scientific use. J Phys Chem Ref Data 31:387–535
Wagner W, Cooper JR, Dittmann A, Kijima J, Kretzschmar HJ, Kruse A, Mareš R, Oguchi K, Sato H, Stocker I, Sifner O, Takaishi Y, Tanishita I, Trübenbach J, Willkommen T (2000) The IAPWS industrial formulation 1997 for the thermodynamic properties of water and steam. J Eng Gas Turbines Power 122:150–182
Wang L, Yang Y, Dong C, Morosuk T, Tsatsaronis G (2014) Systematic optimization of the design of steam cycles using MINLP and differential evolution. J Energy Resour Technol 136:031601
Wang L, Voll P, Lampe M, Yang Y, Bardow A (2015) Superstructure-free synthesis and optimization of thermal power plants. Energy 91:700–711
Wang L, Lampe M, Voll P, Yang Y, Bardow A (2016) Multi-objective superstructure-free synthesis and optimization of thermal power plants. Energy 116:1104–1116

Wang L, Yang Z, Sharma S, Mian A, Lin TE, Tsatsaronis G, Maréchal F, Yang Y (2019) A review of evaluation, optimization and synthesis of energy systems: methodology and application to thermal power plants. Energies 12:73

Zebian H, Gazzino M, Mitsos A (2012) Multi-variable optimization of pressurized oxy-coal combustion. Energy 38:37–57