Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Catalyst acceleration for coordinate descent applied to aircraft assembly simulation

Roman Krasnikov

krasnikov.ra@gmail.com

Peter the Great St. Petersburg Polytechnic University

Maria Titova

Peter the Great St. Petersburg Polytechnic University

Stanislav Baklanov

Peter the Great St. Petersburg Polytechnic University

Sergey Lupuleac

Peter the Great St. Petersburg Polytechnic University

Julia Shinder

Peter the Great St. Petersburg Polytechnic University

Research Article

Keywords: Catalyst acceleration · Coordinate descent · Quadratic programming · Aircraft assembly · Fuselage section mating

Posted Date: November 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10932116/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Optimization and Engineering manuscript No.
(will be inserted by the editor)

A Catalyst acceleration for coordinate descent applied to aircraft assembly simulation

Roman Krasnikov · Maria Titova · Stanislav Baklanov · Sergey Lupuleac · Julia Shinder

Received: date / Accepted: date

Abstract Although a variety of QP solvers are applicable to quadratic programming problems arising in aircraft assembly simulation, computation time – for both problem-specific and general-purpose solvers – remains a significant bottleneck. Addressing this issue, the paper considers a first-order solver based on Catalyst acceleration for coordinate descent. The Catalyst acceleration algorithm and the formulation of a coordinate-descent inner solver for the targeted proximal subproblems in the context of assembly simulation are stated. An analysis of the coordinate descent convergence and complexity is conducted. The invariance of coordinate descent with respect to diagonal scaling has led to a modification of Nesterov's a priori functional gap. A new approach to step control for Catalyst acceleration based on online zero-order minimization of the overall Catalyst complexity is developed. Numerical ex-

R. Krasnikov
Institute of Physics and Mechanics, Peter the Great St. Petersburg Polytechnic University
Polytechnicheskaya, 29, St. Petersburg, 195251, Russia
E-mail: krasnikov.ra@gmail.com

M. Titova
Institute of Physics and Mechanics, Peter the Great St. Petersburg Polytechnic University
Polytechnicheskaya, 29, St. Petersburg, 195251, Russia
E-mail: stefanova.m@list.ru

S. Baklanov
Institute of Physics and Mechanics, Peter the Great St. Petersburg Polytechnic University
Polytechnicheskaya, 29, St. Petersburg, 195251, Russia
E-mail: baklanov.stas@gmail.com

S. Lupuleac
Institute of Physics and Mechanics, Peter the Great St. Petersburg Polytechnic University
Polytechnicheskaya, 29, St. Petersburg, 195251, Russia
E-mail: lupuleac@mail.ru

J. Shinder
Institute of Physics and Mechanics, Peter the Great St. Petersburg Polytechnic University
Polytechnicheskaya, 29, St. Petersburg, 195251, Russia
E-mail: julia_shinder@mail.ru

periments with medium-sized and large-sized assembly models of real aircraft joints are conducted and presented in comparison to other types of QP solvers. Quadratic scalability with respect to problem dimension on a series of fuselage panel assembly models with an increasing number of stringers is experimentally shown. The results of solving contact problems within a full-scale fuselage section mating model confirm the high performance of the presented solver for large-sized contact problems with a bounded condition number.

Keywords Catalyst acceleration · Coordinate descent · Quadratic programming · Aircraft assembly · Fuselage section mating

1 Introduction

Convex quadratic programming (QP) is widely employed across diverse modern computational domains, including machine learning, portfolio optimization, control systems, and structural engineering. As the scale of problems and real-time requirements intensify, research directions focus on the development of algorithms that can handle large-scale problems under time and precision limits. Among the practically significant areas where QP solvers constitute a critical performance bottleneck is the domain of assembly technologies (Sadeghi Tabar et al., 2026) – particularly in the context of high-precision manufacturing, such as aircraft assembly (Lupuleac et al., 2010, 2011). In these applications, QP problems arise naturally from the need to model contact interactions and compute nodal displacements, thereby dictating the overall efficiency and feasibility of the entire simulation workflow.

The underlying analysis typically relies on variation simulation combined with Monte Carlo methods to capture such effects as contact interaction, deformation, friction, dynamic loads, and sealant impact. Each simulation is computationally intensive due to the nonlinear contact problem, which is reformulated as a convex quadratic programming problem using variational principles and finite element discretization. As a result, accurate modeling of assembly processes requires numerous simulation runs, each involving solving a QP problem. For moderate- and large-sized problems, the overall solution time becomes a major limiting factor for process engineers in practice. To address this issue, several optimization methods have been adapted for the assembly problem in recent years, including the Goldfarb-Idnani active-set method (ASM) (Petukhova et al., 2014), Newton projection method (NPM) (Baklanov et al., 2020), and interior-point method (IPM) (Stefanova et al., 2017). The specific structure of the problem, characterized by a fully populated and ill-conditioned Hessian matrix, the necessity of solving numerous QPs with closely varying data, and the moderate-to-large sizes, makes standard general-purpose solvers inefficient in terms of computational time. The distinct structure of the methods brings particular advantages and disadvantages, such as warm-starting, different complexity bounds, sensitivity to the number of active constraints at the solution, different memory footprints, and the need for data preprocessing. As a result, the choice of the most efficient

solver for a given assembly model is not obvious and may vary across different models and configurations. In this context, first-order projection methods emerge as a particularly promising class of QP algorithms that, despite demonstrating notable efficiency in other application areas, have not yet been fully embraced within the domain of compliant assembly simulation. The methods listed above exhibit a worst-case time complexity of at least $\mathcal{O}\left(n^3\right)$, with n denoting the dimension of the problem (Nemirovsky and Yudin, 1983; Nesterov, 2018). The classical results (Nemirovsky and Yudin, 1983; Nesterov, 2018) motivate a time complexity of $\mathcal{O}\left(n^2\sqrt{L/\mu}\right)$, with L/μ denoting the condition number of the Hessian, for the first-order solvers. If the ratio L/μ grows slower than $\mathcal{O}\left(n^2\right)$, rapid convergence on large scale problems can be expected.

The accelerated first-order methods for convex optimization form a modern and rapidly evolving class of methods. Such methods are also referred to as optimal (Nesterov, 2018) because their complexity matches the exact lower bound for first-order algorithms, namely $\mathcal{O}\left(\sqrt{L/\mu}\right)$ oracle calls. Originally, the fast gradient method was presented by Nesterov (1983) for solving unconstrained convex optimization problems with line search for step size selection. This method was generalized by Nesterov for constrained problems (Nesterov, 2018). The optimal methods include other important generalizations, such as the 1-memory and ∞-memory methods for general Bregman functions (Tseng, 2008). The variant of the optimal method with an inexact first-order oracle was widely studied by Devolder et al. (2014). Much research has been conducted on parameter tuning for optimal methods, such as the one-dimensional optimization of the acceleration coefficient proposed by Nesterov et al. (2021) and the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) by Beck and Teboulle (2009). This method incorporates a backtracking line search for an unknown Lipschitz smoothness parameter, as analyzed by Park (2022), and is widely used in applications (Giselsson et al., 2013). Other studies include the Accelerated Hybrid Proximal Extragradient (A-HPE) framework by Monteiro and Svaiter (2013), which serves as a foundational framework for efficient realizations of inexact proximal methods for composite convex optimization, studied by many researchers (Gasnikov et al., 2018, 2019; Lin and Jordan, 2021; Jiang et al., 2021; Marques Alves, 2022; Kolossoski and Monteiro, 2016; Song et al., 2021; Ivanova et al., 2021). An approach using a fully implicit composite formulation within first-order base solvers for an inexact proximal operator calculation is often referred to as Catalyst acceleration, which was studied by Lin et al. (2017) and d'Aspremont et al. (2021).

The optimization techniques considered in this paper include the Catalyst acceleration (Lin et al., 2017; d'Aspremont et al., 2021), coordinate descent (Nesterov, 2012), and a newly proposed online zero-order optimization technique for adaptive step size selection for Catalyst acceleration. Important points for the efficient implementation of the solver, which is based on the Catalyst acceleration approach, include the derivation of a base first-order solver for solving proximal subproblems and an effective step size control strategy.

A coordinate descent is implemented for solving proximal subproblems, which serves as a first-order solver for handling proximal inexactness. The formulation of the method is derived based on the exact one-dimensional minimization of the objective along the specified coordinate. This approach leads to an algorithm that, on one hand, is a general coordinate descent variant as analyzed by Nesterov (2012) and, on the other hand, can be referred to as a nonlinear Gauss-Seidel method following Grippo and Sciandrone (2000). In this paper, Nesterov's a priori error estimate for coordinate descent is modified, specifically in its application to the considered QP problems, to a form independent of the diagonal scaling of the Hessian. This is achieved by introducing a spectral characteristic of the Hessian diagonality. For step size control, an adaptive strategy based on the online zero-order optimization of the overall Catalyst complexity is introduced. It addresses the issue that a priori estimates generally overstate the true functional gap, resulting in an overly conservative (underestimated) optimal Catalyst step size.

The paper is organized as follows. In Section 2, the contact problem formalization within a QP framework is discussed. In Section 3, the solver based on the Catalyst acceleration of coordinate descent is derived in detail, and its convergence and complexity are discussed. In Section 4, a number of numerical experiments are conducted, including a comparative analysis of dual and relative formulations, comparisons to other first-order solvers, as well as comparisons to other types of QP methods for assembly problems, investigations into solver scalability and sensitivity to problem stiffness, and performance testing on a large-scale assembly model of fuselage section mating. The conclusion summarizes the main results regarding the new solver, i.e., its time complexity and efficiency in solving contact problems. Proofs of theorems on the coordinate descent convergence are collected in Appendix A. Technical details and auxiliary lemmas on coordinate descent complexity, modified specifically for the problems under consideration, are collected in Appendix B. Finally, Appendix C provides details on the equation for the optimal step size for Catalyst acceleration within the proposed solver architecture.

2 Problem statement

2.1 Contact problem in the form of quadratic programming problem

The contact problem is a constrained minimization problem in structural mechanics, where the unknown variables are the absolute displacements of nodes in the junction area, total elastic energy is minimized, and constraints prevent the penetration of the contact surfaces.

After reducing the stiffness matrix (Lupuleac et al., 2011), the contact problem can be stated as the following Primal QP problem:

$$\begin{aligned} &\min_{\boldsymbol{x}} \frac{1}{2}\boldsymbol{x}^\top \mathbf{K}\boldsymbol{x} - \boldsymbol{f}^\top \boldsymbol{x}, \\ &\text{s.t. } \mathbf{A}^\top \boldsymbol{x} \leq \boldsymbol{g}, \end{aligned} \tag{1}$$

where

- $\mathbf{K} \in \mathbb{R}^{n \times n}$ – symmetric and positive definite reduced stiffness matrix, block-diagonal with dense blocks,
- $\boldsymbol{f} \in \mathbb{R}^n$ – vector of forces, which models the effects of installed fasteners on assembled parts,
- $\mathbf{A} \in \mathbb{R}^{n \times m}$, $m < n$ – matrix that defines the contact pairs, i.e., $x_i - x_j \leq g_k$,
- $\boldsymbol{g} \in \mathbb{R}^m$ – vector of initial gaps in m contact pairs,
- $\boldsymbol{x} \in \mathbb{R}^n$ – vector of unknown absolute displacements of n computational nodes.

The $\mathbf{K}$ matrix is often ill-conditioned with dense blocks, the matrix A is sparse, and the dimensions n and m are moderate or large (from 1,000 up to 100,000).

2.2 Dual and Relative statements

The property of a trivially calculable orthogonal projection onto the feasible set is critically important for projection-based solvers, which include the NPM and the Catalyst-accelerated coordinate descent method described in this paper. As the matrix A makes it nontrivial to compute the orthogonal projection without a change of variables, the focus of this paper is primarily on two alternative formulations of the contact problem. The Dual QP problem (Lupuleac et al., 2013) is stated as follows:

$$\begin{aligned} \max_{\boldsymbol{\lambda}} & -\frac{1}{2}\boldsymbol{\lambda}^\top \mathbf{Q}\boldsymbol{\lambda} + \boldsymbol{p}^\top \boldsymbol{\lambda} + s, \\ & \text{s.t. } \boldsymbol{\lambda} \geq \mathbf{0}, \end{aligned} \tag{2}$$

where

- $\mathbf{Q} = \mathbf{A}^\top \mathbf{K}^{-1} \mathbf{A} \in \mathbb{R}^{m \times m}$ – symmetric and positive definite dense matrix,
- $\boldsymbol{p} = \mathbf{A}^\top \mathbf{K}^{-1} \boldsymbol{f} - \boldsymbol{g} \in \mathbb{R}^m$,
- $s = -\frac{1}{2}\boldsymbol{f}^\top \mathbf{K}^{-1} \boldsymbol{f} \in \mathbb{R}$,
- $\boldsymbol{\lambda} \in \mathbb{R}^m$ – vector of unknown normal stresses in the contact pairs.

The Dual QP problem has a lower dimension, $m < n$, and box-type constraints. So, the projection $\mathcal{P}_S$ onto the convex closed non-empty feasible set S has the closed coordinate-wise form of

$$\left(\mathcal{P}_S(\lambda)\right)_i = \max\{0, \lambda_i\}. \tag{3}$$

The Relative QP problem (Stefanova et al., 2020) is a variant of the Primal formulated in terms of the relative displacements $\boldsymbol{u} = \mathbf{A}^\top \boldsymbol{x} \in \mathbb{R}^m$:

$$\begin{aligned} \min_{\boldsymbol{u}} & \frac{1}{2}\boldsymbol{u}^\top \tilde{\mathbf{K}} \boldsymbol{u} - \tilde{\boldsymbol{f}}^\top \boldsymbol{u}, \\ & \text{s.t. } \boldsymbol{u} \leq \boldsymbol{g}, \end{aligned} \tag{4}$$

where

- $\tilde{\mathbf{K}} = \left(\mathbf{A}^\top \mathbf{K}^{-1} \mathbf{A}\right)^{-1} \in \mathbb{R}^{m \times m}$ – symmetric and positive definite dense matrix,
- $\tilde{\boldsymbol{f}} = \tilde{\mathbf{K}} \mathbf{A}^\top \mathbf{K}^{-1} \boldsymbol{f}$,
- $\boldsymbol{u} \in \mathbb{R}^m$ – vector of unknown relative displacements.

2.3 Key specifics of problem for solver derivation

Regardless of the formulation (Primal, Dual, or Relative), the Hessian matrix of the QP problem is obtained by reducing the finite element model of the assembly — including its geometry, material properties, and fixtures on the jig — to the junction area where non-penetration conditions define the computational nodes. Unlike conventional structural optimization problems where the system topology changes, assembly analysis features a fixed design, meaning that the Hessian matrix remains constant throughout the study. Meanwhile, part deviations must be incorporated via initial gaps that are inherently uncertain and modeled as a stochastic field, yielding a so-called "cloud of gaps" $\{\boldsymbol{g}_k\}_{k=1}^N$ derived from assembly-line measurements or tolerance analysis, as shown by Lupuleac et al. (2025). The primary engineering objective is to optimize the assembly technology by determining appropriate parameters for temporary fasteners and their optimal installation sequence (Pogarskaia et al., 2019). Consequently, evaluating these technological scenarios requires solving thousands or millions of identical QP problems across various combinations of forces $\boldsymbol{f}$ and gaps $\boldsymbol{g}$, while keeping the Hessian matrix and the matrix of constraints invariant. In this many-query context, the QP solver should be regarded as a highly efficient parametric mapping $(\boldsymbol{f}, \boldsymbol{g}) \mapsto \boldsymbol{x}$, which justifies the computational cost of preprocessing stages, such as constructing the Dual or Relative Hessians.

Many sources on first-order descent methods for convex optimization operate with μ-strongly convex objectives with L-Lipschitz gradient (Nesterov, 2018; Lin et al., 2017). All three QP formulations minimize a quadratic function with a symmetric and positive definite Hessian $\mathbf{Q}$. Hence, $\mu = \lambda_{\min}(\mathbf{Q})$ and $L = \lambda_{\max}(\mathbf{Q})$ follow directly. It will be stated further that a variant of Catalyst acceleration exploits the strong convexity of the objective. Therefore, the parameter μ can be precomputed during preprocessing.

3 First-order QP solver

Classical results on oracle complexity (Nemirovsky and Yudin, 1983) show that first-order methods require $\mathcal{O}\left(\sqrt{L/\mu}\log\left(1/\varepsilon\right)\right)$ oracle calls to achieve a functional gap of $F(\boldsymbol{x}_k) - F(\boldsymbol{x}_*) < \varepsilon$. Nesterov (2013) later proved that the same type of bound – specifically $\mathcal{O}\left(\sqrt{L/\mu}\log\left(1/\varepsilon\right)\right)$ oracle calls with the transfer of $\frac{\mu}{2}\|\boldsymbol{x} - \boldsymbol{x}_0\|^2$ from the smooth part to the non-smooth part of the composite optimization problem statement – holds for proximal methods. In

addition, the same complexity is shown by Monteiro and Svaiter (2013) for the accelerated hybrid proximal extragradient (A-HPE) method within composite optimization and by Lin et al. (2017) for the Catalyst acceleration of the base first-order methods.

Consider a general QP problem of the form:

$$\begin{aligned} \min_{\boldsymbol{x}} F(\boldsymbol{x}) &= \frac{1}{2}\boldsymbol{x}^\top \mathbf{Q}\boldsymbol{x} - \boldsymbol{p}^\top \boldsymbol{x}, \\ &\text{s.t. } \boldsymbol{a} \le \boldsymbol{x} \le \boldsymbol{b}, \end{aligned} \tag{5}$$

with symmetric positive definite $\mathbf{Q} = \{q_{ij}\} \in \mathbb{R}^{m\times m}$, linear term $\boldsymbol{p} \in \mathbb{R}^m$, and box constraints defined by the lower and upper limits $\boldsymbol{a} \in \mathbb{R}^m$ and $\boldsymbol{b} \in \mathbb{R}^m$, respectively. Let $S = \prod_{i=1}^{m} \left[a^i, b^i\right]$ denote the feasible set, and let $\boldsymbol{x}_* = \arg\min_{\boldsymbol{x}\in S} F(\boldsymbol{x})$ be the solution to the problem.

3.1 Catalyst acceleration

This paper focuses on an A-HPE method presented by Monteiro and Svaiter (2013). Aimed at minimizing a proper closed convex function $H : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$, this method has been developed in recent papers, including the strongly convex case and higher-order applications by Marques Alves (2022), the optimal control theory view of A-HPE by Lin and Jordan (2021), and others (Gasnikov et al., 2019, 2018; Song et al., 2021; Jiang et al., 2021; Kolossoski and Monteiro, 2016; Ivanova et al., 2021). The basic problem statement in such studies is a problem of composite optimization, namely,

$$\min_{\boldsymbol{x}\in\mathbb{R}^m} F(\boldsymbol{x}) + G(\boldsymbol{x}) \tag{6}$$

with the convex (or even strongly convex) and smooth (with Lipschitz continuous gradient) function F and the proper closed convex function G. The main part of such methods is the calculation of the prox-operator, defined by

$$\operatorname{prox}_G(\boldsymbol{y}) = \arg\min_{\boldsymbol{x}\in\mathbb{R}^m} \left(G(\boldsymbol{x}) + \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{y}\|^2 \right), \tag{7}$$

where $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^m$. Note that the restrictions imposed on the function G guarantee the correctness of this definition.

In this paper, the fully implicit composite formulation is used, defining the smooth part as identically zero and the non-smooth part as the whole composite objective $H(\boldsymbol{x}) = F(\boldsymbol{x}) + G(\boldsymbol{x})$. Such an approach within the calculation of the inexact proximal operator by some base first-order method is often referred to as *Catalyst acceleration*. This necessitates solving the so-called proximal subproblems

$$\boldsymbol{x}_{k+1} \approx_\delta \operatorname{prox}_{h_k H}(\boldsymbol{y}_k), \tag{8}$$

where $\approx_\delta$ denotes the relative error control (d'Aspremont et al., 2021), defined by the following condition for $\boldsymbol{x}_{k+1}$ and for some $\boldsymbol{g}_H(\boldsymbol{x}_{k+1}) \in \partial H(\boldsymbol{x}_{k+1})$

$$\|h_k \boldsymbol{g}_H(\boldsymbol{x}_{k+1}) + \boldsymbol{x}_{k+1} - \boldsymbol{y}_k\| \leq \delta \|\boldsymbol{x}_{k+1} - \boldsymbol{y}_k\|. \tag{9}$$

Consider a modification of the basic Catalyst acceleration, exploiting a μ-strong convexity of the composite objective (d'Aspremont et al., 2021, Algorithm 24). Its pseudocode is given in Algorithm 1.

Algorithm 1 Catalyst acceleration scheme for optimizing proper closed μ-strongly convex function H

Require: Proper closed μ-strongly convex function H, initial guess $\boldsymbol{x}_0$
  $\boldsymbol{z}_0 \leftarrow \boldsymbol{x}_0$, $A_0 \leftarrow 0$
  for $k = 0, \ldots$ **do**
    Pick the step size $h_k > 0$
    $A_{k+1} \leftarrow A_k + \frac{h_k + 2A_k h_k \mu + \sqrt{4A_k^2 h_k \mu (h_k \mu + 1) + 4A_k h_k (h_k \mu + 1) + h_k^2}}{2}$
    $\boldsymbol{y}_k \leftarrow \boldsymbol{x}_k + \frac{(A_{k+1} - A_k)(A_k \mu + 1)}{A_{k+1} + 2\mu A_{k+1} A_k - \mu A_k^2} (\boldsymbol{z}_k - \boldsymbol{x}_k)$
    Solve proximal subproblem $\boldsymbol{x}_{k+1} \approx_\delta \operatorname{prox}_{h_k H}(\boldsymbol{y}_k)$ with $\delta = \sqrt{1 + h_k \mu}$
    $\boldsymbol{z}_{k+1} \leftarrow \boldsymbol{z}_k + \mu \frac{A_{k+1} - A_k}{1 + \mu A_{k+1}} (\boldsymbol{x}_{k+1} - \boldsymbol{z}_k) - \frac{A_{k+1} - A_k}{1 + \mu A_{k+1}} \boldsymbol{g}_H(\boldsymbol{x}_{k+1})$
  end for

Proximal Subproblem (8) within the formulation of Problem (5) reads as:

$$\begin{aligned} \min_{\boldsymbol{x}} \Phi(\boldsymbol{x}; h_k, \boldsymbol{y}_k) = \frac{1}{2} \boldsymbol{x}^\top \tilde{\mathbf{Q}}(h_k) \boldsymbol{x} - \tilde{\boldsymbol{p}}(h_k, \boldsymbol{y}_k)^\top \boldsymbol{x}, \\ \text{s.t. } \boldsymbol{a} \leq \boldsymbol{x} \leq \boldsymbol{b}, \end{aligned} \tag{10}$$

with the notation

$$\tilde{\mathbf{Q}}(h_k) = \mathbf{Q} + \frac{1}{h_k} \mathbf{I}, \quad \tilde{\boldsymbol{p}}(h_k, \boldsymbol{y}_k) = \boldsymbol{p} + \frac{1}{h_k} \boldsymbol{y}_k. \tag{11}$$

Note that the condition number of $\tilde{\mathbf{Q}}(h)$ reads as $(hL + 1)/(h\mu + 1)$. So, a smaller step size h yields a well-conditioned proximal subproblem, ensuring rapid convergence. However, the overall outer Catalyst acceleration can slow down. In contrast, a large step size speeds up Catalyst convergence but increases the condition number of the proximal subproblem, posing a challenge for the base inner first-order solver.

The use of the $\mathcal{O}(L/\mu)$-method, i.e., the Goldstein-Levitin-Polyak gradient projection method (Levitin and Polyak, 1966), implies an overall Catalyst complexity of $\mathcal{O}\left(\sqrt{L/\mu}\right)$, as stated by Lin et al. (2017) and d'Aspremont et al. (2021). This fact highlights a powerful property of Catalyst acceleration: applying a base first-order solver – fine-tuned for specific problems – to proximal subproblems yields an accelerated solver. Note that the use of the accelerated $\mathcal{O}\left(\sqrt{L/\mu}\right)$-method for solving proximal subproblems yields an

overall Catalyst complexity that is no better than $\mathcal{O}\left(\sqrt{L/\mu}\right)$. This means that the use of the accelerated solver for solving proximal subproblems does not improve the overall Catalyst complexity. Therefore, based on these considerations, a non-accelerated coordinate descent is considered further as the proximal subproblem solver.

In this way, there are two critical aspects in the development of the Catalyst acceleration solver: the base solver for proximal subproblems and, as d'Aspremont et al. (2021) states, a strategy for step size selection.

3.2 Coordinate descent

The proximal subproblem within the Catalyst acceleration framework has the same form as Problem (5). Thus, it is used further for the simplicity of notation.

The coordinate descent consists of sequential coordinate updates

$$x_{k+1}^i = \mathcal{P}_{[a^i,b^i]}\left(x_k^i - \frac{1}{q_{ii}}\nabla_i F(\boldsymbol{x}_k)\right), \tag{12}$$

where $\mathcal{P}_{[a^i,b^i]}$ is the one-dimensional orthogonal projector onto $\left[a^i, b^i\right]$ and $\nabla_i = \partial/\partial x^i$ is the partial derivative. The scheme for derivation of (12) is given in more detail with (15) and (16). The coordinate descent algorithm is presented in pseudocode form as Algorithm 2.

Algorithm 2 Pseudocode of the coordinate descent algorithm

Require: Quadratic programming Problem (5), initial guess $\boldsymbol{x}_0$
  Calculate $\nabla F(\boldsymbol{x}_0)$
  for $k = 0, \ldots$ **do**
    Pick index $i \in \{1, \ldots, m\}$
    $x_{k+1}^i \leftarrow \mathcal{P}_{[a^i,b^i]}\left(x_k^i - \frac{1}{q_{ii}}\nabla_i F(\boldsymbol{x}_k)\right)$
    $\boldsymbol{x}_{k+1} \leftarrow \left[x_k^1, x_k^2, \ldots, x_{k+1}^i, \ldots, x_k^m\right]^\top$
    $\nabla F(\boldsymbol{x}_{k+1}) \leftarrow \nabla F(\boldsymbol{x}_k) + \left(x_{k+1}^i - x_k^i\right)\boldsymbol{q}_i$
  end for

Note that an update of the gradient by the $\nabla F(\boldsymbol{x}_{k+1}) = \nabla F(\boldsymbol{x}_k) + \left(x_{k+1}^i - x_k^i\right)\boldsymbol{q}_i$, where $\boldsymbol{q}_i$ is the i-th column of $\mathbf{Q}$, takes only $\mathcal{O}(m)$ operations, which is preferred to a naive gradient calculation, whose time complexity is $\mathcal{O}\left(m^2\right)$.

In the implementation of the proposed coordinate descent a greedy index selection strategy is used, namely

$$i_{k+1} = \arg\max_{i \in \{1,\ldots,m\}} \left|\left(\nabla_{\mathcal{P}} F(\mathbf{x}_k)\right)_i\right|, \tag{13}$$

where $\nabla_{\mathcal{P}}$ is a subgradient of F with the least euclidean norm, given by

$$\left(\nabla_{\mathcal{P}} F(\boldsymbol{x})\right)_i = \begin{cases} 0, \ x^i = a^i, \nabla_i F(\boldsymbol{x}) > 0 \text{ or } x^i = b^i, \nabla_i F(\boldsymbol{x}) < 0, \\ \nabla_i F(\boldsymbol{x}), \ \text{otherwise.} \end{cases} \tag{14}$$

Such strategy is targeted on vanishing the largest subgradient component, which contributes to the fastest reduction of the proximal subproblem stationary condition, namely the left-hand side of (9).

3.2.1 Convergence and complexity analysis

For convergence analysis, following Nesterov (2012), one-dimensional gradient mapping $\xi_S^i(\boldsymbol{x})$ is introduced along with its vectorization $\boldsymbol{r}_S^i(\boldsymbol{x})$

$$\begin{aligned} \xi_S^i(\boldsymbol{x}) &= \arg \min_{y \in [a^i, b^i]} \left\{ F(\boldsymbol{x}) + \nabla_i F(\boldsymbol{x}) \left(y - x^i\right) + \frac{q_{ii}}{2} \left(y - x^i\right)^2 \right\} = \\ &= \mathcal{P}_{[a^i, b^i]} \left(x^i - \frac{1}{q_{ii}} \nabla_i F(\boldsymbol{x}) \right), \\ \boldsymbol{r}_S^i(\boldsymbol{x}) &= [x^1, x^2, \ldots, \xi_S^i(\boldsymbol{x}), \ldots, x^m]^\top, \end{aligned} \tag{15}$$

where $\nabla_i = \partial / \partial x^i$. Gradient mapping $\xi_S^i(\boldsymbol{x})$ is the optimal value of the variable i (in the sense of the optimization of the local quadratic majorant), and $\boldsymbol{r}_S^i(\boldsymbol{x})$ is the vector obtained by substituting $\xi_S^i(\boldsymbol{x})$ in place of the i-th coordinate. Note that the closed form for $\xi_S^i(\boldsymbol{x})$ follows directly from rearranging

$$\begin{aligned} \nabla_i F(\boldsymbol{x}) \left(y - x^i\right) + \frac{q_{ii}}{2} \left(y - x^i\right)^2 = \\ = \frac{q_{ii}}{2} \left(y - x^i + \frac{1}{q_{ii}} \nabla_i F(\boldsymbol{x}) \right)^2 - \frac{1}{2 q_{ii}} \left(\nabla_i F(\boldsymbol{x}) \right)^2 . \end{aligned} \tag{16}$$

Also, for every $y \in \mathbb{R}$ and $i \in \{1, \ldots, m\}$

$$F\left(\boldsymbol{x} + \left(y - x^i\right) \boldsymbol{e}_i\right) = F(\boldsymbol{x}) + \nabla_i F(\boldsymbol{x}) \left(y - x^i\right) + \frac{q_{ii}}{2} \left(y - x^i\right)^2, \tag{17}$$

where $\boldsymbol{e}_i$ is the i-th standard basis vector. It makes it possible to alternatively define $\xi_S^i(\boldsymbol{x})$ as the exact one-dimensional minimizer

$$\xi_S^i(\boldsymbol{x}) = \arg \min_{y \in [a^i, b^i]} \left\{ F\left(\boldsymbol{x} + \left(y - x^i\right) \boldsymbol{e}_i\right) \right\}. \tag{18}$$

This leads to the duality of the considered coordinate descent with Gauss-Seidel methods (Grippo and Sciandrone, 2000), which is shown rigorously (see Theorem 2) and discussed further.

The generalization of the objective partial derivative, a reduced partial derivative $\gamma_S^i(\boldsymbol{x})$, is given by

$$\gamma_S^i(\boldsymbol{x}) = q_{ii} \left(x^i - \xi_S^i(\boldsymbol{x})\right). \tag{19}$$

Note that $a^i = -\infty$ and $b^i = +\infty$ yield $\gamma_S^i(\boldsymbol{x}) = \nabla_i F(\boldsymbol{x})$.

Theorem 1 formulates the quantitative characteristic of the descent with $\xi_S^i(\boldsymbol{x})$ step (see Appendix A for proofs and details).

Theorem 1 *Let* $F(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top \mathbf{Q}\boldsymbol{x} - \boldsymbol{p}^\top \boldsymbol{x}$ *with symmetric positive definite matrix* $\mathbf{Q}$*. Then for any* $\boldsymbol{x} \in S$

$$F(\boldsymbol{x}) - F\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right) \geq \frac{1}{2q_{ii}}\gamma_S^i(\boldsymbol{x})^2. \tag{20}$$

Corollary 1 *For every* $i \in \{1, \ldots, m\}$ *the equalities* $\gamma_S^i(\boldsymbol{x}_*) = 0$ *and* $\xi_S^i(\boldsymbol{x}_*) = x_*^i$ *hold. On the contrary, if for every* $i \in \{1, \ldots, m\}$ *it holds that* $\gamma_S^i(\boldsymbol{x}) = 0$*, then* $\boldsymbol{x} = \boldsymbol{x}_*$*.*

Formula (15) uses the quadratic majorant of F, which yields a gradient step with a one-dimensional Lipschitz constant. The present paper exploits the fact that F is quadratic. As the next theorem shows, under this approach, the reduced partial derivative γ_S^i vanishes after a single one-dimensional gradient descent step.

Theorem 2 *Under the assumptions of Theorem 1, for any* $\boldsymbol{x} \in \mathbb{R}^m$

$$\gamma_S^i\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right) = 0. \tag{21}$$

The presented coordinate descent method for solving the quadratic programming problem with box constraints can be viewed from two complementary angles. First, it can be interpreted as the Gauss-Seidel method, which is based on the exact minimization of the objective function along the chosen coordinate. Second, Expression (15) can be viewed as a one-dimensional gradient mapping, as considered by Nesterov (2012) for general coordinate descent methods with an arbitrary partition of the $\mathbb{R}^m$ space. Such a duality is not a general property of coordinate descent; rather, it arises specifically due to the structure of the quadratic programming problem with box constraints under consideration. As demonstrated below, exploiting the exact form of the objective function and constraints can yield a modification of the Nesterov general coordinate descent a priori error estimate (see Appendix B for details).

Theorem 3 *Let* $\{\boldsymbol{x}_k\}_{k=0}^{\infty}$ *be the sequence generated by the coordinate descent applied to Problem* (5)*, and let the next index of the coordinate to update be chosen randomly and distributed uniformly on* $\{1, \ldots, m\}$*. Then, for every* $k \geq 0$*,*

$$\mathbb{E}\left[F\left(\boldsymbol{x}_k\right) - F\left(\boldsymbol{x}_*\right)\right] \leq \left(1 + \frac{1}{\hat{\mu}}\right)\left(1 - \frac{2\hat{\mu}}{m(1+\hat{\mu})}\right)^k \left[F\left(\boldsymbol{x}_0\right) - F\left(\boldsymbol{x}_*\right)\right], \tag{22}$$

where $\hat{\mu} = \lambda_{min}(\hat{\mathbf{Q}})$ *is the minimal eigenvalue of the scaled Hessian* $\hat{\mathbf{Q}}$ *consisting of elements* $\hat{q}_{ij} = q_{ij}/\sqrt{q_{ii}q_{jj}}$*.*

Theorem 3 gives the modified a priori error estimate, which is based on Nesterov's result for general coordinate descent and is applied to a scaled problem that preserves the original trajectory of the coordinate descent. After simplifying $2\hat{\mu}/(1+\hat{\mu}) \geq \hat{\mu}$, this yields a complexity of coordinate descent that

reads as $\mathcal{O}\left((m/\hat{\mu})\log(1/\varepsilon)\right)$ iterations in terms of reaching the functional gap $F(\boldsymbol{x}_k) - F(\boldsymbol{x}_*) < \varepsilon$.

The parameter $\hat{\mu} \in (0, 1]$ is a quantitative characteristic of the spectral distance between the scaled Hessian $\hat{\mathbf{Q}}$ and the identity $\mathbf{I}$, namely, $1 - \hat{\mu} \leq \|\mathbf{I} - \hat{\mathbf{Q}}\|_2$. Thus, $\hat{\mu} = 1$ is the best possible option for diagonal Hessian $\mathbf{Q} = \operatorname{diag}\{q_{11}, q_{22}, \ldots, q_{mm}\}$ and $\hat{\mu}$, which is close to zero, indicates strong correlation between variables, which decreases the convergence rate of the coordinate descent under consideration.

3.2.2 Application to solving proximal subproblems and overall Catalyst complexity

Consider a priori error estimate for method $\mathcal{M}$ of solving proximal subproblems

$$\mathbb{E}\left[\Phi(\boldsymbol{x}_k; h, \boldsymbol{y}) - \Phi(\boldsymbol{x}_*; h, \boldsymbol{y})\right] \leq \\ \leq C_{\mathcal{M},h}\left(1 - \tau_{\mathcal{M},h}\right)^k \left[\Phi(\boldsymbol{x}_0; h, \boldsymbol{y}) - \Phi(\boldsymbol{x}_*; h, \boldsymbol{y})\right], \quad (23)$$

where $\boldsymbol{x}_* = \arg\min_{\boldsymbol{x} \in S} \Phi(\boldsymbol{x}; h, \boldsymbol{y})$ is the precise solution of the proximal subproblem. It is shown by Lin et al. (2017) that, for given constants $C_{\mathcal{M},h}$ and $\tau_{\mathcal{M},h}$ one can construct an expression for the overall Catalyst acceleration complexity. Theoretically, such an approach is based on the minimization of

$$\min_{h \geq 0} \frac{1}{\tau_{\mathcal{M},h}} \sqrt{\frac{h\mu + 1}{h\mu}}, \quad (24)$$

which, asymptotically, is the Catalyst acceleration complexity (for a constant step $h_k \equiv h$), i.e., the number of iterations of method $\mathcal{M}$ needed to reach the desired functional gap (neglecting the ε term). After some analysis (see Appendix C for details), the equation for the optimal constant step size h_*, i.e., the minimizer for (24), reads as

$$\left(L_{\mathbf{D}}\hat{\mu} + 2(\hat{\mu} - 1)\mu\right) h_*^2 + (3\hat{\mu} - 1)h_* + \frac{1}{L_{\mathbf{D}}} = 0, \quad (25)$$

where $L_{\mathbf{D}} = \max_{i \in \{1,\ldots,m\}} q_{ii}$. Although the final expression for h_* can be very large for the theoretical analysis of the overall Catalyst acceleration complexity, in the worst case when $\hat{\mu} = \mu/L_{\mathbf{D}}$, the optimal step size is $h_* = 1/(L_{\mathbf{D}} - 2\mu)$, and the overall complexity of the Catalyst acceleration is $\mathcal{O}\left(m\sqrt{L_{\mathbf{D}}/\mu}\log(1/\varepsilon)\right)$ iterations of coordinate descent.

Many numerical experiments with different assembly models have shown that the theoretically optimal step underestimates the practical one. This is a consequence of the tendency of a priori estimates to significantly overestimate the real functional gap, which is natural for almost all a priori estimates. Moreover, the worst-case value $\hat{\mu} = \mu/L$ itself underestimates the practically optimal step size. This observation motivates an advanced step size selection

strategy based on using the number of iterations needed by $\mathcal{M}$ to solve the proximal subproblem, substituting this number into the expression for the Catalyst acceleration complexity, and minimizing this complexity online by varying the step size at each iteration.

3.3 Step size selection strategy

With an a priori error estimate given by d'Aspremont et al. (2021, Corollary 5.8), it can be shown that the Catalyst complexity (number of iterations of the outer algorithm) is $\mathcal{O}\left(\sqrt{\frac{1+\mu h}{\mu h}}\log\left(\frac{1}{\varepsilon}\right)\right)$. Therefore, with $B_{\mathcal{M},h}$ being an upper bound for the number of iterations of the inner solver for the proximal subproblem with step size h, the overall complexity is $\mathcal{O}\left(B_{\mathcal{M},h}\sqrt{\frac{1+\mu h}{\mu h}}\log\left(\frac{1}{\varepsilon}\right)\right)$. The derivation of the criterion (24) for the theoretically optimal constant step size $h_k \equiv h$ is based on the fact that asymptotically $B_{\mathcal{M},h} \approx 1/\tau_{\mathcal{M},h}$. This is stated by d'Aspremont et al. (2021) for deterministic methods of non-constrained optimization and by Lin et al. (2017) for both deterministic and stochastic methods.

As discussed before, such a theoretical estimate for the optimal step size is based on a priori error estimates, which often significantly overestimate the real functional gap. Therefore, an adaptive online step size control is presented in this paper. It is based on the number of inner iterations $N_{\text{inner}}(h_k, \boldsymbol{y}_k)$, which is known after the k-th iteration of the Catalyst scheme, namely

$$\min_{h_k \geq 0} \psi(h_k) = N_{\text{inner}}(h_k, \boldsymbol{y}_k)\sqrt{\frac{1+h_k\mu}{h_k\mu}}. \tag{26}$$

The following numerical study was conducted to develop a heuristic way to optimize ψ criterion online. On the test assembly models, a specific contact problem was solved within the constant step size approach, varying the step size for different runs. Then the ratio of the number of all iterations of the coordinate descent to the number of Catalyst iterations is considered an estimate for the average N_{inner}. Experimentally obtained curves $\psi(h)$ were then constructed. The characteristic form of $\psi(h)$ constructed within the dual formulation of one test contact problem is given in Figure 1.

Empirically, the criterion $\psi(h)$ appears to be unimodal, slightly resembling a parabola on logarithmic axes. This observation motivates the following Algorithm 3 for online zero-order minimization of such a criterion within the variation of the logarithmic step size increment.

Consider the beginning of the algorithm ($k = 0$) with $h_0 = 1/(L_{\mathbf{D}} - 2\mu)$ denoting the theoretically optimal step size. After performing the Catalyst acceleration iteration, the values of N_{inner} and $\psi(h_k)$ can be calculated. If $\psi(h_k) < \psi(h_{k-1})$, then the step size should be increased, which is achieved by multiplying it by the factor ν or, equivalently, by increasing $\log h$ by $\log \nu$. This process continues unless $\psi(h_k) > \psi(h_{k-1})$, which means that the criterion ψ

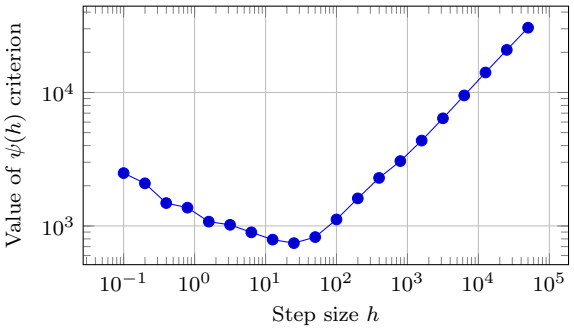


Fig. 1: Experimentally obtained characteristic form of $\psi(h)$ criterion

Algorithm 3 General scheme of the adaptive step size selection within the Catalyst acceleration

Pick $h_0 = \frac{1}{L_{\mathbf{D}}-2\mu}$, $\nu_0 > 1$, $\nu_{\min} > 1$
for $k = 0, \dots$ **do**
  Perform an outer iteration of the Catalyst acceleration ▷ See Alg. 1
  if $\psi(h_k) \geq \psi(h_{k-1})$ **then**
    if $\max\left\{\sqrt{\nu_k}, \sqrt{\frac{1}{\nu_k}}\right\} < \nu_{\min}$ **then**
      $\nu_{k+1} \leftarrow 1/\nu_k$
    else
      $\nu_{k+1} \leftarrow 1/\sqrt{\nu_k}$
    end if
  else
    $\nu_{k+1} \leftarrow \nu_k$
  end if
  Update the step size $h_{k+1} \leftarrow \nu_{k+1} h_k$
end for

Remark. The value of $\psi(h_{-1})$ is considered $+\infty$ for simplicity of notation.

has started to increase, and this needs to be prevented by multiplying the step size by $1/\nu$, thus decreasing $\log h$ by $\log \nu$. A strategy for varying ν is also provided for a rapid search of the optimal step neighborhood. It is followed by refinement through more precise step size variation. The value of $|\log \nu|$ is approximately bounded below by $\log \nu_{\min}$ to make the algorithm responsive to stochastic fluctuations of N_{inner}.

After fine-tuning on the test assembly models, the values $\nu_0 = 2$ and $\nu_{\min} = 1.15$ were chosen as the most effective hyperparameters for the developed heuristic algorithm.

4 Numerical results

In this section, numerical results are presented for the solver under consideration as applied to the assembly models of real aircraft joints. An important advantage of the Dual formulation in comparison with the Relative one

is shown. The numerical efficiency of the presented solver is evaluated on assembly problems, comparing it against other first-order solvers as well as against other types of QP methods adapted earlier for such problems, including the Goldfarb-Idnani active-set method (ASM), the Newton projection method (NPM), and the interior-point method (IPM). An analysis of the solver scalability and sensitivity to problem stiffness is conducted. Finally, the simulation of the joint of full-scale cylindrical fuselage sections is performed using the developed numerical approach. At the same time, the execution time of the algorithm on such highly complex problems enables multiple multi-variant computations. Thus, rigorous engineering analysis becomes possible even for this, arguably the most challenging procedure in the commercial aircraft assembly process.

All numerical experiments reported in this section were performed in single-thread mode using the ASRP (Assembly Simulation of Riveting Process) software (Lupuleac et al., 2010).

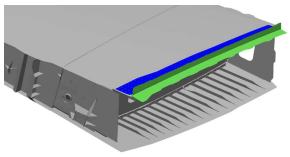
4.1 Analysis of the efficiency of exploiting structural properties of considered QP problems

Six structural models are selected to evaluate the proposed method: two industrial cases of wing-to-fuselage joint (representing the lower and upper joints; see (Lupuleac et al., 2019c) for details), a partial joint of two ring-shaped surfaces, an assembly of fuselage panels (Zaitseva et al., 2025), a pylon assembly (Lupuleac et al., 2019b), and the attachment of the reinforcement splice to the fuselage section (Lupuleac et al., 2019a). The geometry and junction areas of these models are shown in Figure 2. The third model (Figure 2c) was specifically designed to ensure that, for a given dimension, its characteristics — particularly the condition number — fall within a specified range. In this model, the junction area spans only one-tenth of the total circumference (shown in blue in Figure 2c), and the contact problem is solved exclusively in this region. Physically, this implies that the interpenetration of parts is allowed outside this zone.

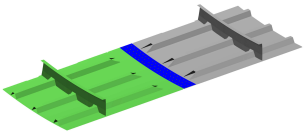
The quantitative characteristics of the medium-sized assembly models are given by Table 1. Dual and Relative formulations have the same condition number L/μ, as their Hessians are mutually inverse. In contrast, the Dual formulation has lower complexity parameters $L_{\mathbf{D}}/\mu$ and $1/\hat{\mu}$, which are important characteristics in terms of the theoretical complexity analysis of the first-order solver. This observation motivates the hypothesis that the developed first-order solver is more efficient in the Dual formulation of the contact problems.

There are two other optimal methods considered in the following numerical experiments. These are given with the simple iterative scheme

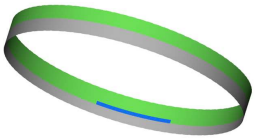
$$\begin{aligned} \boldsymbol{x}_{k+1} &= \mathcal{P}_S\left(\boldsymbol{y}_k - \frac{1}{L}\nabla F(\boldsymbol{y}_k)\right), \\ \boldsymbol{y}_{k+1} &= \boldsymbol{x}_{k+1} + \beta_{k+1}\left(\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\right). \end{aligned}$$



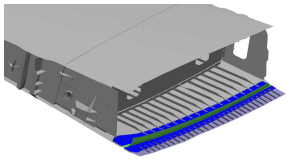
(a) Upper wing-to-fuselage joint (Lupuleac et al., 2019c)



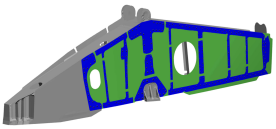
(b) Assembly of small fuselage panels (Zaitseva et al., 2025)



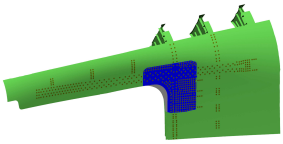
(c) A partial joint of two ring-shaped surfaces



(d) Lower wing-to-fuselage joint (Lupuleac et al., 2019c)



(e) Pylon assembly (Lupuleac et al., 2019b)



(f) Attaching a reinforcing splice to a fuselage section (Lupuleac et al., 2019a)

Fig. 2: Medium-sized assembly models

Table 1: Quantitative characteristics of medium-sized assembly models

Model		(a)	(b)	(c)	(d)	(e)	(f)
Number of nodes n		5206	4522	6766	3890	7554	11143
Number of constraints m		2603	2261	3383	2476	3777	6493
Dual	$\log_{10}(L/\mu)$	5.41	5.47	5.98	7.88	7.76	7.71
	$\log_{10}(L_{\mathbf{D}}/\mu)$	3.04	3.16	3.45	5.90	4.92	5.44
	$\log_{10}(1/\hat{\mu})$	2.66	2.89	3.34	5.23	4.82	4.80
Relative	$\log_{10}(L/\mu)$	5.41	5.47	5.98	7.88	7.76	7.71
	$\log_{10}(L_{\mathbf{D}}/\mu)$	5.23	5.16	5.55	7.34	7.46	7.59
	$\log_{10}(1/\hat{\mu})$	4.89	5.05	5.43	6.90	6.84	6.89

The one with $\beta_k = \left(\sqrt{L} - \sqrt{\mu}\right) / \left(\sqrt{L} + \sqrt{\mu}\right)$ is referred to as "Nesterov" addressing the famous Constant Step Scheme by Nesterov (2018). This method, applied to a μ-strongly convex objective with a L-Lipschitz gradient, has linear convergence and an iteration complexity of $\mathcal{O}\left(\sqrt{L/\mu}\log(1/\varepsilon)\right)$. The second one uses $\beta_k = k/(k+3)$, as proposed by Tseng (2008). This variant is referred to as "Tseng" and, when applied to convex objectives with a Lipschitz gradient, has a sublinear convergence rate of $\mathcal{O}\left(1/k^2\right)$, i.e., the functional gap has an upper bound of the form C/k^2, where C is some constant (dependent on the initial guess and problem properties) and k is the number of iterations. Therefore, "Tseng" solver can be viewed as the most general one, while "Nesterov" exploits the strong convexity of the objective. These methods are the most widely used in applications and thus are compared to the proposed Catalyst-accelerated coordinate descent, which is referred to as "A-HPE" (addressing the original naming by Monteiro and Svaiter).

Numerical experiments show a practical advantage in exploiting the structural properties of the QP problems under consideration for constructing a proper solver. The results of the considered numerical experiments are given in Figure 3. The x-axis is set up as the number of active constraints of the Primal QP problem. The variation in the number of active constraints is achieved by varying the number of installed fasteners with the uniform initial gap $\boldsymbol{g}$ of 0.6 mm. The more fasteners are placed, the more contacts arise, and thus the number of active constraints increases. Note that the exact mapping of installed fasteners to the number of active constraints is nonlinear and highly dependent on the specific assembly model. Plots for models (d), (e) and (f) omit the lines for "Tseng" as the solving time for these models reaches 10^5 seconds per QP problem.

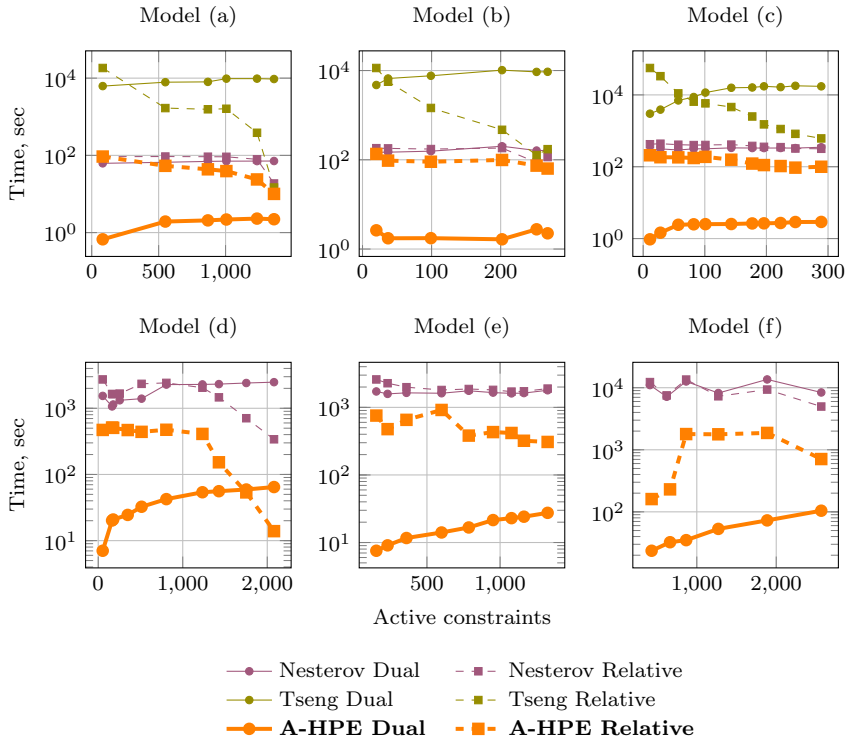


Fig. 3: Comparison of optimal first-order methods

The numerical experiments have shown that for almost every specific contact problem, the first-order solver is faster by a factor of 10 to 100 in the Dual formulations than in the Relative one. Note that all plots show the opposite trends for Dual and Relative formulations. This is explained by the initial guess, which is taken as $\mathbf{0}$ in Dual problems and $\boldsymbol{g}$ in Relative ones. Thus, a

lower number of active constraints means that the Dual initial guess is closer to the exact solution, while a high number of active constraints has the opposite effect. Despite this fact, for example, for model (a), the developed solver is still faster by a factor of approximately 5 to 10 on problems with a number of active constraints up to 1500, which corresponds to 58% of all possible constraints m. Note that for all assembly models "Tseng" solver is the slowest one, and the A-HPE is the fastest, which is a direct consequence of exploiting structural properties of the considered QP problems. Moreover, for models (a), (b) and (c), the difference between "A-HPE Relative" and "Nesterov" (both Dual and Relative formulation) is insignificant – this can be explained by the proximity of parameters $L_{\mathbf{D}}/\mu$ (and $1/\hat{\mu}$) to L/μ (see Table 1). Also, the modified complexity parameters $L_{\mathbf{D}}/\mu$ and $1/\hat{\mu}$ are lower for Dual formulations, which leads to a faster convergence of the A-HPE applied to Dual problems. These results confirm that exploiting problem-specific structure leads to practical acceleration beyond what is predicted by worst-case complexity bounds alone.

4.2 Solver performance dependence on problem dimension and stiffness

In the following experiments, it is natural to consider models (a), (b), and (c) as models with a medium condition number, while models (d), (e), and (f) are considered ill-conditioned, having a higher condition number L/μ and complexity parameters $L_{\mathbf{D}}/\mu$ and $1/\hat{\mu}$. The dual dimension m in these types of models varies, making it possible to observe the effect of an increase in dimension for a fixed problem stiffness. Thus, the models considered can be used as a simple preliminary experiment to study the scalability and sensitivity of the first-order solver to problem stiffness.

The results of solving contact problems for medium-sized models using different solvers are shown in Figure 4. The following solvers were compared:

- ASM – the Goldfarb–Idnani active set method,
- IPM – the primal–dual interior-point method,
- NPM – the Newton projection method.

The A-HPE solver outperforms the other solvers in models (a), (b), (c), and (f) in the Dual formulation. Nevertheless, it performs poorly in comparison to NPM and ASM (mostly in the Relative formulation) in models (d) and (e). These are the two models with the lowest dual dimension m and high QP problem stiffness, so the poor performance of A-HPE on them correlates with the theoretical complexity estimates. At the same time, A-HPE slightly outperforms the other solvers on the ill-conditioned model (f). This is explained by its high dual dimension m: the $\mathcal{O}(m^2)$ time complexity of A-HPE (for fixed $L_{\mathbf{D}}/\mu$, neglecting the ε term) is better than that of the other solvers, whose time complexity is at least $\mathcal{O}(m^3)$ due to solving the linear systems.

Table 2 contains the mean solver time averaged over different fastener configurations (different numbers of active constraints) for each of the six

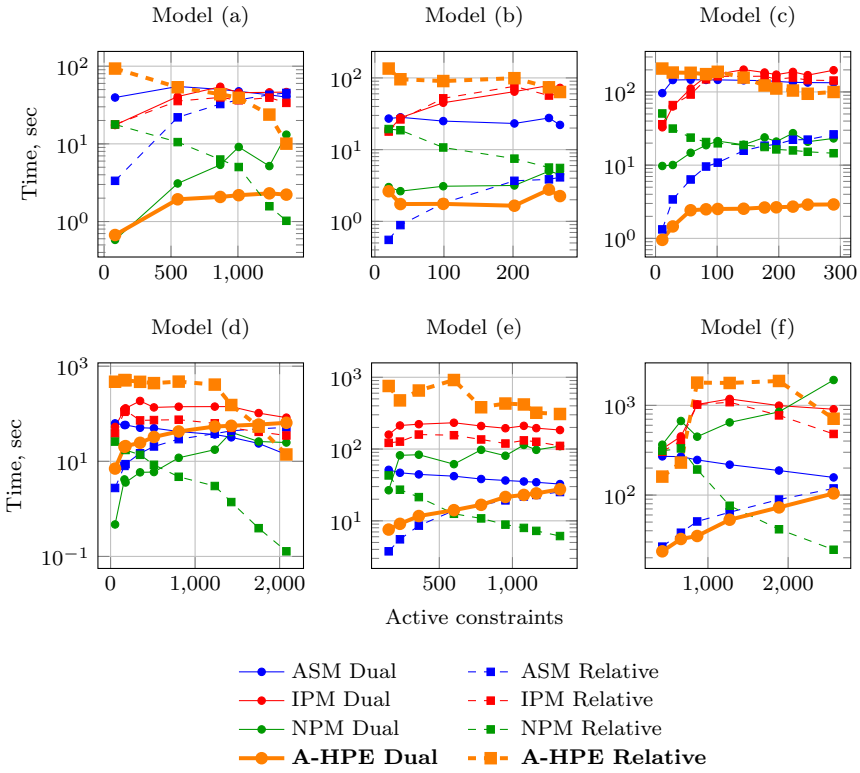


Fig. 4: Solver time for the medium-sized assembly models with different solvers

medium-sized assembly models, the solvers and formulations considered. The first-order A-HPE solver (for Dual formulations) outperforms other solvers by a factor of approximately 3 in model (a) and 6 in model (c). However, the A-HPE for Relative formulations performs poorly, in fact being the slowest solver, which is a consequence of the higher complexity parameters $L_{\mathbf{D}}/\mu$ and $1/\hat{\mu}$ discussed before. The A-HPE also underperforms in comparison with the NPM for the Relative formulation in model (d). This shows that the A-HPE solver scales well, in comparison to other solvers, with the increase of problem dimension while condition number remains approximately constant. In contrast, increasing the problem stiffness leads to poor performance, as the experiment with model (d) has shown.

Generally, the second-order NPM solver is typically the fastest among all other solvers except A-HPE in the numerical experiments conducted. Therefore, it is of particular interest to compare these two solvers. This comparison, shown in Figure 5, was performed on models used in real-world engineering analysis of various aircraft assembly joints. The dot color in the figure indicates which solver is faster on average for each specific model. This comparison con-

Table 2: Mean solver time (seconds) for the medium-sized assembly models

Solver	Formulation	(a)	(b)	(c)	(d)	(e)	(f)
ASM	Dual	45.80	25.55	137.16	42.74	40.15	224.66
	Relative	30.47	2.48	14.16	26.56	15.29	64.58
IPM	Dual	41.76	51.09	150.65	123.70	202.29	812.98
	Relative	33.75	48.31	129.74	67.64	132.06	677.49
NPM	Dual	6.08	3.57	19.08	14.11	83.66	814.87
	Relative	7.03	11.23	22.30	9.51	16.11	162.89
A-HPE	Dual	1.90	2.13	2.38	38.09	17.23	53.54
	Relative	43.89	92.98	147.94	350.18	517.93	1089.62

firms the important property of the developed A-HPE solver: it outperforms the NPM on the large-sized models that are not too ill-conditioned.

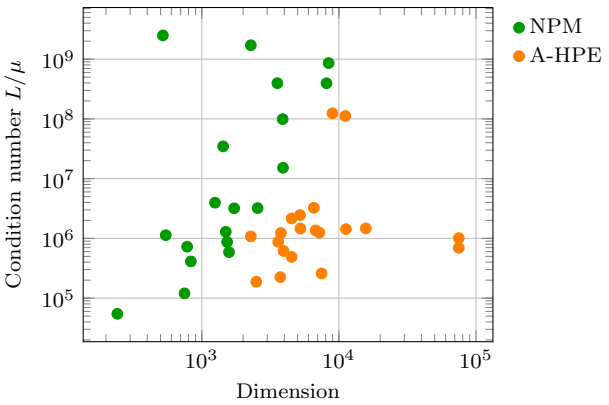


Fig. 5: Performance comparison of A-HPE and NPM on different assembly models

Note that the practically relevant models shown in Figure 5 have $\log_{10}(L/\mu)$ in the range of approximately 6 to 7 and span a wide range of Primal dimensions n from 500 to 15000. This means that increasing the problem dimension in the assembly modeling technology considered nearly does not affect the problem stiffness. Considering such a set of assembly models, $\mathcal{O}(m^2)$ scalability can be expected as the dimension m of the Dual problems increases. In the next section, an important set of fuselage section assembly models with extended junction area is studied, which experimentally confirms this fact.

4.3 Scalability in fuselage assembly models

To evaluate the scalability of the proposed method, several configurations of the fuselage section joint are investigated (see Figure 6). A finite element model consisting of two cylindrical fuselage sections is adopted as the baseline configuration (Zaitseva et al., 2025). The computational complexity is varied systematically by adjusting the spatial extent of the junction area in the respective contact problems. In the first model, the junction spans only 1/90 of the total circumference and encompasses a single stringer. The subsequent configurations are scaled up to cover 1/45 of the circumference (corresponding to 2 stringers), followed by larger zones containing 5, 9 and 15 stringers, respectively. Similar to the partial joint model of two ring-shaped surfaces (Figure 2c), this formulation implies that the mutual penetration of the components is physically permissible outside the designated junction zone.

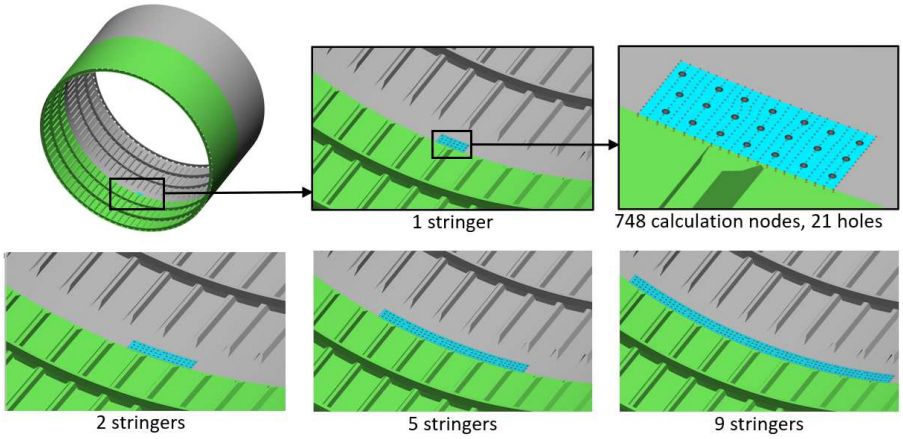


Fig. 6: Assembly models of cylindrical fuselage sections

The results of the scalability experiment with an increasing number of stringers in the fuselage assembly model are presented in Table 3. The possible logarithmic slope χ in the time complexity estimation for the solvers $\mathcal{O}(m^{\chi})$ is determined numerically. The values $\chi_i = \log(t_i/t_{i-1}) / \log(m_i/m_{i-1})$, where t_i denotes the solution time (in seconds), are calculated and presented in Table 3. The table reports the solution time and the logarithmic slope χ for the Dual problems with approximately 50% of the installed fasteners. It shows that the A-HPE has better scalability, with approximately $\mathcal{O}\left(m^2\right)$ time complexity. Although its performance is comparable to the NPM for 1-5 stringers, on the large-sized models, A-HPE significantly outperforms the other solvers. On the model with 15 stringers the advantage is by a factor of up to 7 compared with NPM.

Table 3: Experimentally obtained time and logarithmic slope χ

Number of stringers		1	2	5	9	15
Primal dimension n		782	1530	3774	6766	11254
Dual dimension m		391	765	1887	3383	5627
Time, sec	ASM	0.10	0.93	17.09	100.78	488.46
	IPM	0.45	4.25	52.46	250.48	1020.75
	NPM	0.03	0.28	4.11	29.46	109.94
	A-HPE	0.10	0.34	1.75	5.34	15.31
Log-slope χ	ASM	—	3.28	3.23	3.04	3.09
	IPM	—	3.36	2.78	2.68	2.75
	NPM	—	3.23	2.98	3.37	2.75
	A-HPE	—	1.90	1.81	1.91	2.06

The data from the experiment are shown graphically in Figure 7. It visually shows that A-HPE has a lower logarithmic slope χ compared to the other solvers. In addition, the logarithmic slopes χ of ASM, IPM, and NPM oscillate between 2.5 and 3.5, while that of A-HPE is less volatile and approximately equal to 2. This observation experimentally confirms the $\mathcal{O}\left(m^2\right)$ time complexity of the new solver on the specified fuselage assembly models.

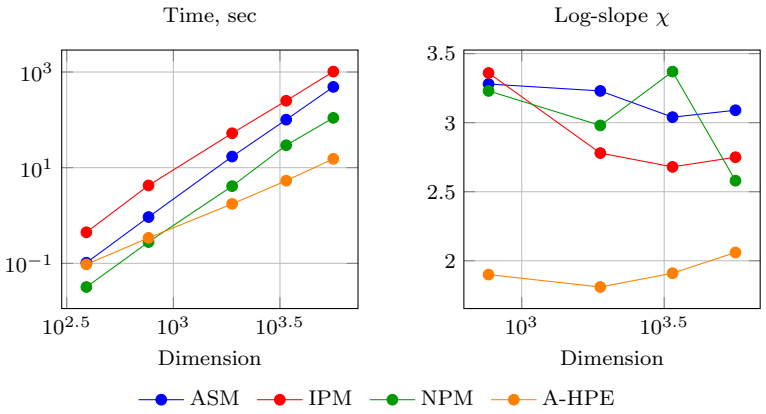


Fig. 7: Plots of Table 3 data for models of fuselage sections assembly

In conclusion, the first-order A-HPE solver exhibits $\mathcal{O}(m^2)$ scalability in the series of fuselage assembly models considered. Simple extrapolation motivates the expectation of high performance on large-scale assembly models. This will be examined further using the example of joining two cylindrical fuselage sections.

4.4 Computational efficiency on a full-scale fuselage barrel joint model

The mating of composite barrel-shaped fuselage sections is one of the most challenging operations in the assembly process of modern passenger aircraft, such as the Boeing 787 (Gates, 2007) or the Airbus A350. This underscores the need for accurate modeling of the assembly process. Numerous studies have addressed fuselage section assembly modeling, ranging from partial-circumference panels (Lupuleac et al., 2011; Qu et al., 2026; Xue et al., 2026) to complete cylindrical barrel joints (Wang et al., 2023; Cheng et al., 2015; Wen et al., 2018, 2019; Du et al., 2019, 2022; Lutz et al., 2022). However, the large dimensions of the junction area, the topological complexity and mechanical flexibility of the structure, as well as its high sensitivity to variations in external conditions (e.g., loads applied by fasteners or deviations in fixture positioning on the assembly jig), require computational models of considerable complexity to adequately represent the assembly process. Furthermore, as noted above, the engineering simulation of assembly processes requires accounting for all possible geometric deviations of every structural component, which inevitably leads to numerous multi-scenario analyses. These factors pose a major challenge for the development of the underlying mathematical framework. To make numerical tools for assembly process simulation practically useful for engineering applications, it is necessary to develop solvers capable of producing highly accurate solutions (since the cost of prediction errors in this field is extremely high) while maintaining high computational efficiency. In practice, the acceptable solution time for a single simulation is on the order of seconds or minutes rather than hours or days. To the best of the authors' knowledge, until recently this challenge had not been fully addressed, and all existing models represented a compromise between:

- The mechanical complexity of the model (whether or not contact interactions between structural components are taken into account);
- The resolution of the computational mesh (whether or not established engineering meshing standards are satisfied);
- The topology of the modeled joint (the complete cylindrical junction or only a local segment).

The proposed algorithm enables the construction of an assembly model free from the compromises outlined above. Specifically, it is sufficiently detailed and mechanically comprehensive to reproduce all major physical phenomena observed during the assembly process while maintaining a computation time of only a few minutes (if wet assembly is not considered, as was done by Eliseev et al. (2024a,b)).

The joint of two full-scale fuselage sections with a diameter of 6 m is considered (see Fig. 8). The fragment of the finite element mesh is presented on the left-hand side. The residual gap distribution, calculated for a uniform initial gap of 0.6 mm with 5% of the fasteners installed, is shown on the right-hand side of Figure 8.

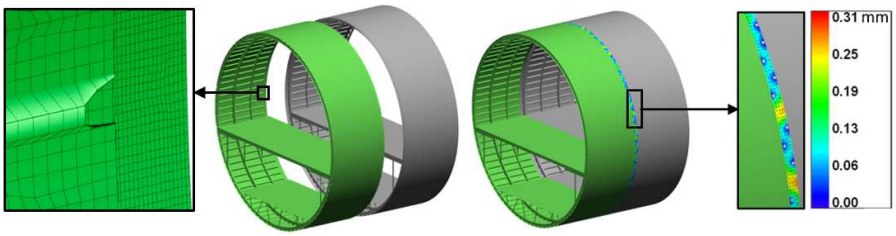


Fig. 8: Full-scale fuselage barrel joint model

The results of the numerical experiment with the full-scale fuselage barrel joint model ($n = 74880$, $m = 37440$) are presented in Figure 9. Numerical experiments utilize two different constant gaps between the sections (physically corresponding to the connection of two cylindrical surfaces with different radii): 0.6 mm (default) and 6 mm (large). The results show significantly better performance of A-HPE compared to NPM. It is a consequence of the $\mathcal{O}\left(m^2\right)$ scalability of the A-HPE solver. The trend of increasing solver time as the number of active constraints of the Primal QP problem increases holds for both solvers because the initial guess is the one with all constraints active (namely, zero for the Dual formulation). While the fastener configuration is fixed for both gaps, the results for a larger gap of 6 mm show fewer active constraints of the Primal QP problem. This is because the contact is less likely to appear with a larger gap between the parts being assembled.

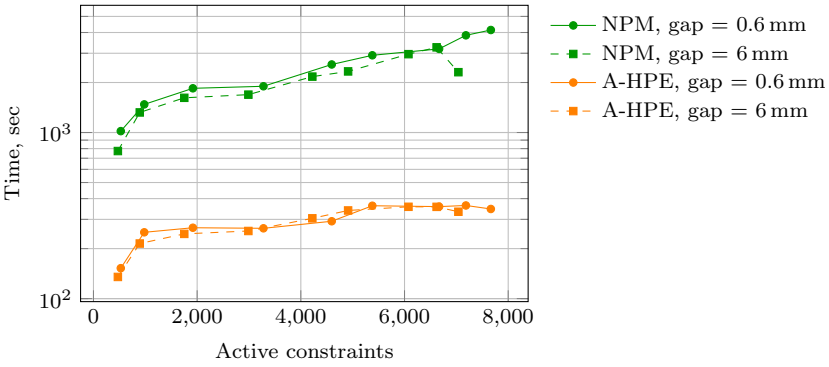


Fig. 9: Solver time comparison for NPM and A-HPE on the fuselage barrel joint model

The numerical data from the experiment are presented in Table 4. It shows an A-HPE speedup of about 6 to 12 compared to NPM. Note that the significant difference in the initial gap value does not have much influence on the solver time. The performance of A-HPE on the fuselage barrel joint model

makes it possible to simulate the assembly process taking into account deviations of the structural components, as described by Lupuleac et al. (2025).

Table 4: A-HPE speedup over NPM on the fuselage barrel joint model with constant initial gaps

Number of fasteners		100	200	400	600	800	1000	1400	1600	1890
Gap = 0.6 mm										
Active constraints		527	980	1918	3279	4596	5379	6667	7184	7665
Time, sec	NPM	1019	1477	1846	1898	2568	2917	3188	3847	4142
	A-HPE	153	251	267	265	292	362	358	364	347
A-HPE Speedup, ×		6.69	5.89	6.91	7.16	8.79	8.05	8.90	10.58	11.95
Gap = 6 mm										
Active constraints		471	893	1752	2988	4222	4912	6076	6619	7039
Time, sec	NPM	774	1321	1618	1690	2170	2333	2963	3257	2308
	A-HPE	135	215	245	255	304	339	357	357	333
A-HPE speedup, ×		5.73	6.15	6.59	6.62	7.13	6.87	8.30	9.13	6.92

In conclusion, a specific feature of the assembly simulation technology considered, which arises as a boundedness of the problem condition number with increasing dimensions, makes an algorithm with the time complexity of $\mathcal{O}\left(m^2\sqrt{L_{\mathbf{D}}/\mu}\log\left(1/\varepsilon\right)\right)$ a reasonable choice. For example, the condition number L/μ for the fuselage barrel joint model under consideration is equal to $1.69 \cdot 10^5$. Note that the corresponding logarithmic value $\log_{10}(L/\mu) \approx 5.23$ is the lowest in comparison with all medium-sized models (see $\log_{10}(L/\mu)$ values in Table 1).

Although the advantages of the presented solver are widely discussed, some disadvantages should be noted. Firstly, the numerical experiments in Sections 4.1 and 4.2 have shown that the first-order A-HPE solver is highly dependent on the problem stiffness. This is a natural property of all first-order solvers, including the Catalyst acceleration with different non-accelerated first-order solvers for handling proximal inexactness. Nevertheless, some assembly models with high QP problem stiffness can pose a challenge for the presented solver. Secondly, the A-HPE solver has shown the greatest solver time gap between Dual and Relative formulations. Although the reasons for such an effect are well studied in Section 4.1, the solver can perform poorly on dense matrices with $\hat{\mu}$ close to zero. Despite the disadvantages, the presented solver has made a qualitative shift in the solver time for the fuselage barrel joint model, making it possible to conduct a statistical analysis for the validation of the fastener configuration and its optimization. The results of statistical experiments with random gaps $\boldsymbol{g}$, presented in Figure 10, show a nearly linear dependence of the solver time on the number of active constraints of the Primal QP problem.

The statistical experiment covers uniform gaps of 0.6 mm and 6 mm, as well as 10 randomly generated gaps, resulting in a total of 112 contact problems. The black line shows a linear regression with an intercept of 174 seconds, a slope of 25.9 seconds per 1000 active constraints, and an R^2 value of 0.921. Note that no outliers have appeared, confirming the stability of the solver time in the statistical analysis with random initial gaps.

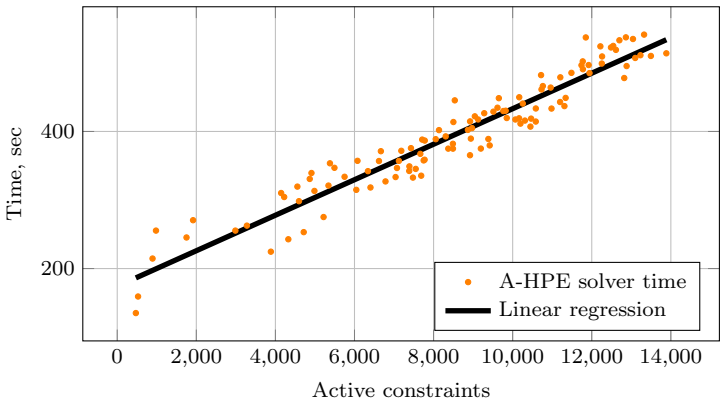


Fig. 10: A-HPE solver time for different fastener configurations and initial gaps

5 Conclusion

This paper presents a first-order solver based on Monteiro and Svaiter's A-HPE framework for solving contact problems in aircraft assembly simulation, which arise as Dual and Relative QP problems. The solver is a Catalyst acceleration for coordinate descent – a fully implicit A-HPE method with coordinate descent handling proximal inexactness. An analysis of coordinate descent convergence and complexity has shown that a spectral characteristic $\hat{\mu}$ of Hessian diagonality is a key parameter that determines the coordinate descent complexity of $\mathcal{O}\left((m/\hat{\mu})\log\left(1/\varepsilon\right)\right)$ iterations to reach the desired functional gap ε. When applied to solving proximal subproblems, the worst-case analysis leads to the overall Catalyst complexity of $\mathcal{O}\left(m\sqrt{L_{\mathbf{D}}/\mu}\log(1/\varepsilon)\right)$ coordinate descent iterations with a time cost of $\mathcal{O}(m)$. A step size selection strategy for Catalyst acceleration, based on online zero-order optimization of the overall Catalyst complexity with respect to the number of inner solver iterations, is presented. The solver performance is validated with a series of numerical experiments on assembly models of real aircraft joints. Quadratic scalability with respect to the problem dimension is shown. In contrast, increasing the problem stiffness leads to poor performance of the first-order solver considered.

High performance on the large-scale fuselage barrel joint model is confirmed. Finally, the solver presented in this paper has high potential for fast calculations on not too ill-conditioned QP problems with box constraints, while high stiffness remains a challenge for the first-order solver.

Acknowledgements The research was supported by the Russian Science Foundation (project No. 22-19-00062-П, `https://rscf.ru/project/22-19-00062/`).

Statements and Declarations

Competing interests

The authors have no competing interests to declare that are relevant to the content of this article.

Funding

The research was supported by the Russian Science Foundation (project No. 22-19-00062-П, `https://rscf.ru/project/22-19-00062/`).

Ethics approval and Consent to participate

Not applicable.

Data availability

Replication of the results of the numerical experiments is supported by the detailed description of the algorithm. The specific model data and numerical experiment results are available from the corresponding author upon reasonable request.

Author contribution

Roman Krasnikov developed the first-order solver and studied its theoretical properties. Maria Titova and Stanislav Baklanov performed the literature review, summarized the relevant findings, and provided details on other solvers mentioned in the manuscript. Sergey Lupuleac described the problem statement, reviewed references on engineering challenges in aircraft assembly simulation, and designed the physical formulation for numerical experiments. Julia Shinder constructed finite element models, conducted comprehensive testing of solvers for the models described in the manuscript, and compiled the numerical results. All authors contributed to the writing and review of the manuscript.

A Convergence analysis of the proposed coordinate descent

Theorem 1 and Corollary 1 propose a quantitative characteristic of descent with $\boldsymbol{r}_S^i(\boldsymbol{x})$ and a corresponding local optimality measure.

Theorem 1 *Let $F(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top \mathbf{Q}\boldsymbol{x} - \boldsymbol{p}^\top \boldsymbol{x}$ with symmetric positive definite matrix $\mathbf{Q}$. Then for any $\boldsymbol{x} \in S$*

$$F(\boldsymbol{x}) - F\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right) \geq \frac{1}{2q_{ii}}\gamma_S^i(\boldsymbol{x})^2. \tag{20}$$

Proof By the definition of ξ_S^i the following variational inequality holds

$$\forall z \in \left[a^i, b^i\right] : \ \nabla_i F\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right)\left(z - \xi_S^i(\boldsymbol{x})\right) \geq 0. \tag{27}$$

Since $\nabla F(\boldsymbol{x}) = \mathbf{Q}\boldsymbol{x} - \boldsymbol{p}$, it follows that

$$\nabla_i F(\boldsymbol{x}) - \nabla_i F\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right) = q_{ii}\left(x^i - \xi_S^i(\boldsymbol{x})\right) = \gamma_S^i(\boldsymbol{x}). \tag{28}$$

Substituting (28) into (27) and setting $z = x^i$ gives

$$\nabla_i F\left(\boldsymbol{x}\right)\left(x^i - \xi_S^i(\boldsymbol{x})\right) \geq \gamma_S^i(\boldsymbol{x})\left(x^i - \xi_S^i(\boldsymbol{x})\right). \tag{29}$$

Exploiting the exact form of F yields

$$\begin{aligned} F\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right) - F(\boldsymbol{x}) &= \nabla_i F\left(\boldsymbol{x}\right)\left(\xi_S^i(\boldsymbol{x}) - x^i\right) + \frac{q_{ii}}{2}\left(\xi_S^i(\boldsymbol{x}) - x^i\right)^2 \\ &\leq -\frac{1}{q_{ii}}\gamma_S^i(\boldsymbol{x})^2 + \frac{1}{2q_{ii}}\gamma_S^i(\boldsymbol{x})^2 = -\frac{1}{2q_{ii}}\gamma_S^i(\boldsymbol{x})^2, \end{aligned} \tag{30}$$

which completes the proof. □

Corollary 1 *For every $i \in \{1, \ldots, m\}$ the equalities $\gamma_S^i\left(\boldsymbol{x}_*\right) = 0$ and $\xi_S^i\left(\boldsymbol{x}_*\right) = x_*^i$ hold. On the contrary, if for every $i \in \{1, \ldots, m\}$ it holds that $\gamma_S^i\left(\boldsymbol{x}\right) = 0$, then $\boldsymbol{x} = \boldsymbol{x}_*$.*

Proof According to (20), using the optimality of $\boldsymbol{x}_*$,

$$0 \geq F\left(\boldsymbol{x}_*\right) - F\left(\boldsymbol{r}_S^i\left(\boldsymbol{x}_*\right)\right) \geq \frac{1}{2q_{ii}}\gamma_S^i\left(\boldsymbol{x}_*\right)^2. \tag{31}$$

Now, consider arbitrary $i \in \{1, \ldots, m\}$ and arbitrary $y \in \left[a^i, b^i\right]$. Since $\gamma_S^i(\boldsymbol{x}) = 0$ implies $\xi_S^i(\boldsymbol{x}) = x^i$, it follows that

$$\nabla_i F(\boldsymbol{x})\left(y - x^i\right) = \nabla_i F\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right)\left(y - \xi_S^i(\boldsymbol{x})\right) \geq 0 \tag{32}$$

by the definition of ξ_S^i. Summing up these inequalities by all $i \in \{1, \ldots, m\}$ proves that $\boldsymbol{x}$ is indeed the minimizer. □

Therefore, $\sum\limits_{i=1}^{m} \gamma_S^i(\boldsymbol{x})^2$ can be considered a measure of the optimality of $\boldsymbol{x}$.

Theorem 2 motivates a similarity of the proposed coordinate descent with Gauss-Seidel methods, which optimize an objective on the subspace of $\mathbb{R}^m$ precisely.

Theorem 2 *Under the assumptions of Theorem 1, for any $\boldsymbol{x} \in \mathbb{R}^m$*

$$\gamma_S^i\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right) = 0. \tag{21}$$

Proof Using the definition of ξ_S^i, the following inequalities hold

$$\begin{aligned} &\nabla_i F\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right)\left(\xi_S^i\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right)-\xi_S^i(\boldsymbol{x})\right)\geq 0\\ &\nabla_i F\left(\boldsymbol{r}_S^i\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right)\right)\left(\xi_S^i(\boldsymbol{x})-\xi_S^i\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right)\right)\geq 0. \end{aligned} \tag{33}$$

The sum of these inequalities gives the following

$$\left(\nabla_i F\left(\boldsymbol{r}_S^i\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right)\right)-\nabla_i F\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right)\right)\left(\xi_S^i\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right)-\xi_S^i(\boldsymbol{x})\right)\leq 0, \tag{34}$$

while in the explicit form of the function F,

$$\begin{aligned} &\left(\nabla_i F\left(\boldsymbol{r}_S^i\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right)\right)-\nabla_i F\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right)\right)\left(\xi_S^i\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right)-\xi_S^i(\boldsymbol{x})\right)\\ &\qquad = q_{ii}\left(\xi_S^i\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right)-\xi_S^i(\boldsymbol{x})\right)^2, \end{aligned} \tag{35}$$

which yields $\gamma_S^i\left(\boldsymbol{r}_S^i(\boldsymbol{x})\right)=0$. □

B Complexity analysis of the proposed coordinate descent

Let $\|\cdot\|_{\mathbf{D}}$, given by

$$\|\boldsymbol{x}\|_{\mathbf{D}}^2=\sum_{i=1}^m q_{ii}\left(x^i\right)^2, \tag{36}$$

be the energy norm induced by the diagonal part $\mathbf{D}=\operatorname{diag}\mathbf{Q}$. The a priori error estimates by Nesterov use the value $R\left(\boldsymbol{x}_0\right)$ given by

$$R\left(\boldsymbol{x}_0\right)=\max_{\boldsymbol{x}\in S}\left\{\max_{\boldsymbol{x}_*\in X_*}\|\boldsymbol{x}-\boldsymbol{x}_*\|_{\mathbf{D}}\mid F\left(\boldsymbol{x}\right)\leq F\left(\boldsymbol{x}_0\right)\right\}, \tag{37}$$

where X_* is the optimal set. The next lemma gives an explicit upper estimate of $R\left(\boldsymbol{x}_0\right)$ for the class of quadratic programming problems under consideration.

Lemma 1 *Let $F(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^\top\mathbf{Q}\boldsymbol{x}-\boldsymbol{p}^\top\boldsymbol{x}$ with a symmetric positive definite Hessian, $\mu=\lambda_{min}(\mathbf{Q})$ be the minimal eigenvalue of $\mathbf{Q}$ and $L_{\mathbf{D}}=\max\limits_{i\in\{1,\ldots,m\}}q_{ii}$ be the greatest diagonal value of $\mathbf{Q}$. Then*

$$\frac{1}{2}R^2(\boldsymbol{x}_0)\leq\frac{L_{\mathbf{D}}}{\mu}\left(F\left(\boldsymbol{x}_0\right)-F\left(\boldsymbol{x}_*\right)\right). \tag{38}$$

Proof The problem considered has a unique solution, $X_*=\{\boldsymbol{x}_*\}$, as it is a problem of minimization of the strongly convex, continuous, and coercive functional F on a non-empty closed convex set S. Note that

$$\boldsymbol{x}^\top\mathbf{Q}\boldsymbol{x}\geq\mu\sum_{i=1}^m\left(x^i\right)^2=\mu\sum_{i=1}^m\frac{1}{q_{ii}}q_{ii}\left(x^i\right)^2\geq\frac{\mu}{L_{\mathbf{D}}}\sum_{i=1}^m q_{ii}\left(x^i\right)^2=\frac{\mu}{L_{\mathbf{D}}}\|\boldsymbol{x}\|_{\mathbf{D}}^2 \tag{39}$$

implies strong convexity of the objective in the norm $\|\cdot\|_{\mathbf{D}}$ with parameter of $\mu/L_{\mathbf{D}}$. Therefore, for every $\boldsymbol{x}\in S$,

$$F(\boldsymbol{x})\geq F(\boldsymbol{x}_*)+\langle\nabla F(\boldsymbol{x}_*),\boldsymbol{x}-\boldsymbol{x}_*\rangle+\frac{\mu}{2L_{\mathbf{D}}}\|\boldsymbol{x}-\boldsymbol{x}_*\|_{\mathbf{D}}^2. \tag{40}$$

The optimality of $\boldsymbol{x}_*$ results in

$$\frac{1}{2}\|\boldsymbol{x}-\boldsymbol{x}_*\|_{\mathbf{D}}^2\leq\frac{L_{\mathbf{D}}}{\mu}\left(F\left(\boldsymbol{x}\right)-F\left(\boldsymbol{x}_*\right)\right). \tag{41}$$

Consider an arbitrary $\boldsymbol{x}\in S$ such that $F(\boldsymbol{x})\leq F(\boldsymbol{x}_0)$; then for every such $\boldsymbol{x}$

$$\frac{1}{2}\|\boldsymbol{x}-\boldsymbol{x}_*\|_{\mathbf{D}}^2\leq\frac{L_{\mathbf{D}}}{\mu}\left(F\left(\boldsymbol{x}_0\right)-F\left(\boldsymbol{x}_*\right)\right). \tag{42}$$

Taking the maximum of the left-hand side completes the proof. □

Nesterov's a priori error estimate for general coordinate descent (Nesterov, 2012, Theorem 5) is stated for the variant of the method with the random index uniformly distributed on $\{1, \ldots, m\}$ and estimates the expectation of the functional gap, denoted further as $\mathbb{E}\left[F\left(\boldsymbol{x}_k\right) - F\left(\boldsymbol{x}_*\right)\right]$. Applied to quadratic programming Problem (5), it reads as follows.

Theorem 4 *Let $\{\boldsymbol{x}_k\}_{k=0}^{\infty}$ be the sequence generated by the coordinate descent applied to Problem* (5) *and let the next index of the coordinate to update be chosen randomly and distributed uniformly on $\{1, \ldots, m\}$. Then, for every $k \geq 0$,*

$$\mathbb{E}\left[F\left(\boldsymbol{x}_k\right) - F\left(\boldsymbol{x}_*\right)\right] \leq \left(1 + \frac{L_{\mathbf{D}}}{\mu}\right)\left(1 - \frac{2\mu}{m(L_{\mathbf{D}} + \mu)}\right)^k \left[F\left(\boldsymbol{x}_0\right) - F\left(\boldsymbol{x}_*\right)\right]. \tag{43}$$

Theorem 4 is based on a priori estimates of the error (in the sense of the functional gap) for a wide range of coordinate descent methods. Intuitively, it is clear that the considered coordinate descent, applied to the quadratic programming problem with box constraints (5), must be equally effective for every problem with a diagonal Hessian $\mathbf{Q}$. Yet, for example, for $\mathbf{Q} = \mathbf{I} = \operatorname{diag}\{1, 1\}$ the value of $\mu / L_{\mathbf{D}}$ is equal to the best possible $\mu / L_{\mathbf{D}} = 1$ and for $\mathbf{Q} = \operatorname{diag}\{1, 100\}$ this value is $\mu / L_{\mathbf{D}} = 0.01$. Thus, such an a priori estimate is sensitive to the diagonal scaling of Problem (5), whereas the coordinate descent, as shown below in explicit form, is not. Therefore, a more precise a priori error estimate can be constructed for Problem (5) based on Nesterov's estimate.

A scaled analog of Problem (5) is introduced as follows. Let

$$\mathbf{D}^{\alpha} = \operatorname{diag}\left\{q_{11}^{\alpha}, q_{22}^{\alpha}, \ldots, q_{mm}^{\alpha}\right\}. \tag{44}$$

Substituting $\hat{\boldsymbol{x}} = \mathbf{D}^{1/2}\boldsymbol{x}$ into Problem (5) yields the equivalent scaled problem

$$\begin{aligned} \min_{\hat{\boldsymbol{x}}} \hat{F}(\hat{\boldsymbol{x}}) &= \frac{1}{2}\hat{\boldsymbol{x}}^{\top}\hat{\mathbf{Q}}\hat{\boldsymbol{x}} - \hat{\boldsymbol{p}}^{\top}\hat{\boldsymbol{x}}, \\ \text{s.t. } \hat{\boldsymbol{a}} &\leq \hat{\boldsymbol{x}} \leq \hat{\boldsymbol{b}}, \end{aligned} \tag{45}$$

where $\hat{\mathbf{Q}} = \mathbf{D}^{-1/2}\mathbf{Q}\mathbf{D}^{-1/2}$, $\hat{\boldsymbol{p}} = \mathbf{D}^{-1/2}\boldsymbol{p}$, $\hat{\boldsymbol{a}} = \mathbf{D}^{1/2}\boldsymbol{a}$, and $\hat{\boldsymbol{b}} = \mathbf{D}^{1/2}\boldsymbol{b}$. It is clear that $\hat{F}(\hat{\boldsymbol{x}}) = F(\boldsymbol{x})$, $\hat{\boldsymbol{x}}_* = \mathbf{D}^{1/2}\boldsymbol{x}_*$, and $\hat{\boldsymbol{x}}$ is feasible in Problem (45) if and only if $\boldsymbol{x}$ is feasible in Problem (5). The next lemma shows the invariance of the coordinate descent trajectory with respect to such a transformation.

Lemma 2 *Let $\{\boldsymbol{x}_k\}_{k=0}^{\infty}$ and $\{\hat{\boldsymbol{x}}_k\}_{k=0}^{\infty}$ be the sequences generated by coordinate descent for problems* (5) *and* (45)*, respectively. Moreover, let $\hat{\boldsymbol{x}}_0 = \mathbf{D}^{1/2}\boldsymbol{x}_0$ and at every iteration the same coordinate index was picked to update. Then, for every $k \geq 0$ trajectories of the coordinate descent are the same: $\hat{\boldsymbol{x}}_k = \mathbf{D}^{1/2}\boldsymbol{x}_k$.*

Proof The proof is based on the important property of the projector onto box constraints: for every $c > 0$ and arbitrary $a, b, x \in \mathbb{R}$

$$\mathcal{P}_{[ca, cb]}(cx) = c\mathcal{P}_{[a,b]}(x), \tag{46}$$

which can be proved directly. The statement is proved by mathematical induction. The initial $\hat{\boldsymbol{x}}_0 = \mathbf{D}^{1/2}\boldsymbol{x}_0$ holds by the theorem statement. Now, under the assumption $\hat{\boldsymbol{x}}_k = \mathbf{D}^{1/2}\boldsymbol{x}_k$, denoting by $\boldsymbol{q}_i$ the i-th row of $\mathbf{Q}$ and $\hat{\boldsymbol{q}}_i$ the i-th row of $\hat{\mathbf{Q}}$, it can be shown that

$$\begin{aligned} \hat{x}_{k+1}^i &= \mathcal{P}_{\left[\hat{a}^i, \hat{b}^i\right]}\left(\hat{x}_k^i - \hat{\boldsymbol{q}}_i^{\top}\hat{\boldsymbol{x}}_k + \hat{p}^i\right) \\ &= \mathcal{P}_{\left[\sqrt{q_{ii}}a^i, \sqrt{q_{ii}}b^i\right]}\left(\sqrt{q_{ii}}x_k^i - \frac{1}{\sqrt{q_{ii}}}\boldsymbol{q}_i^{\top}\mathbf{D}^{-1/2}\mathbf{D}^{1/2}\boldsymbol{x}_k + \frac{1}{\sqrt{q_{ii}}}p^i\right) \\ &= \sqrt{q_{ii}}\mathcal{P}_{\left[a^i, b^i\right]}\left(x_k^i - \frac{1}{q_{ii}}\boldsymbol{q}_i^{\top}\boldsymbol{x}_k + \frac{1}{q_{ii}}p^i\right) = \sqrt{q_{ii}}x_{k+1}^i, \end{aligned} \tag{47}$$

which completes the proof. □

Therefore, solving Problem (5) can be addressed as solving the scaled one (45) in terms of variables $\hat{\boldsymbol{x}}$. Applying Theorem 4 to the scaled Problem (45) under $\hat{F}(\hat{\boldsymbol{x}}_k) = F(\boldsymbol{x}_k)$ and $\hat{F}(\hat{\boldsymbol{x}}_*) = F(\boldsymbol{x}_*)$ yields Theorem 3 from the main part of the paper.

Theorem 3 *Let $\{\boldsymbol{x}_k\}_{k=0}^{\infty}$ be the sequence generated by the coordinate descent applied to Problem (5), and let the next index of the coordinate to update be chosen randomly and distributed uniformly on $\{1, \ldots, m\}$. Then, for every $k \geq 0$,*

$$\mathbb{E}\left[F(\boldsymbol{x}_k) - F(\boldsymbol{x}_*)\right] \leq \left(1 + \frac{1}{\hat{\mu}}\right)\left(1 - \frac{2\hat{\mu}}{m(1+\hat{\mu})}\right)^k \left[F(\boldsymbol{x}_0) - F(\boldsymbol{x}_*)\right], \tag{22}$$

where $\hat{\mu} = \lambda_{min}(\hat{\mathbf{Q}})$ is the minimal eigenvalue of the scaled Hessian $\hat{\mathbf{Q}}$ consisting of elements $\hat{q}_{ij} = q_{ij}/\sqrt{q_{ii}q_{jj}}$.

Returning to the considered example, both $\mathbf{Q} = \mathrm{diag}\{1, 1\}$ and $\mathbf{Q} = \mathrm{diag}\{1, 100\}$ yield $\hat{\mathbf{Q}} = \mathrm{diag}\{1, 1\}$ and the best possible $\hat{\mu} = 1$. Thus, such a modified error estimate is invariant with respect to diagonal scaling, which is an expected property of the considered coordinate descent.

Note that the a priori error estimate given by Theorem 3 is more precise than the basic one provided by Theorem 4, as shown in the next remark.

Remark 1

$$\hat{\mu} \geq \frac{\mu}{L_{\mathbf{D}}}. \tag{48}$$

Proof The minimal eigenvalue is the minimum of the Rayleigh quotient:

$$\hat{\mu} = \min_{\boldsymbol{x}\neq\mathbf{0}} \frac{\boldsymbol{x}^\top \hat{\mathbf{Q}}\boldsymbol{x}}{\boldsymbol{x}^\top\boldsymbol{x}} = \min_{\boldsymbol{x}\neq\mathbf{0}} \frac{\boldsymbol{x}^\top \mathbf{D}^{1/2}\hat{\mathbf{Q}}\mathbf{D}^{1/2}\boldsymbol{x}}{\boldsymbol{x}^\top\mathbf{D}^{1/2}\mathbf{D}^{1/2}\boldsymbol{x}} = \min_{\boldsymbol{x}\neq\mathbf{0}} \frac{\boldsymbol{x}^\top \mathbf{Q}\boldsymbol{x}}{\boldsymbol{x}^\top\mathbf{D}\boldsymbol{x}} \geq \frac{\min\limits_{\boldsymbol{x}\neq\mathbf{0}} \frac{\boldsymbol{x}^\top\mathbf{Q}\boldsymbol{x}}{\boldsymbol{x}^\top\boldsymbol{x}}}{\max\limits_{\boldsymbol{x}\neq\mathbf{0}} \frac{\boldsymbol{x}^\top\mathbf{D}\boldsymbol{x}}{\boldsymbol{x}^\top\boldsymbol{x}}} = \frac{\mu}{L_{\mathbf{D}}}, \tag{49}$$

which completes the proof. □

Finally, the coordinate descent under consideration has $\mathcal{O}\left((m/\hat{\mu})\log(1/\varepsilon)\right)$ iteration complexity and $\mathcal{O}\left(\left(m^2/\hat{\mu}\right)\log(1/\varepsilon)\right)$ time complexity.

C Derivation of the equation for the optimal constant step size for Catalyst acceleration

Let $\hat{\mu}_h$ be defined as

$$\hat{\mu}_h = \lambda_{\min}\left(\mathbf{D}_h^{-1/2}\left(\mathbf{Q} + \frac{1}{h}\mathbf{I}\right)\mathbf{D}_h^{-1/2}\right) \tag{50}$$

where $\mathbf{D}_h^{\alpha} = \mathrm{diag}\left\{(q_{11} + 1/h)^{\alpha}, (q_{22} + 1/h)^{\alpha}, \ldots, (q_{mm} + 1/h)^{\alpha}\right\}$. For the coordinate descent applied to the proximal subproblem, it is derived that $\tau_{\mathcal{M},h} = \hat{\mu}_h/m$. The next lemma provides a reasonable estimate for $\hat{\mu}_h$, making it possible to determine an equation for the optimal constant step h.

Lemma 3

$$\hat{\mu}_h \geq \frac{\hat{\mu}L_{\mathbf{D}}h + 1}{L_{\mathbf{D}}h + 1}. \tag{51}$$

Proof In this proof, let $s = h^{-1}$. Again, using the Rayleigh quotient, it can be shown that

$$\hat{\mu}_h = \min_{\boldsymbol{x}\neq\mathbf{0}} \frac{\boldsymbol{x}^\top\mathbf{Q}\boldsymbol{x} + s\boldsymbol{x}^\top\boldsymbol{x}}{\boldsymbol{x}^\top\mathbf{D}\boldsymbol{x} + s\boldsymbol{x}^\top\boldsymbol{x}} \geq \min_{\boldsymbol{x}\neq\mathbf{0}} \frac{\hat{\mu}\boldsymbol{x}^\top\mathbf{D}\boldsymbol{x} + s\boldsymbol{x}^\top\boldsymbol{x}}{\boldsymbol{x}^\top\mathbf{D}\boldsymbol{x} + s\boldsymbol{x}^\top\boldsymbol{x}}. \tag{52}$$

Consider the function $f(t) = \frac{\hat{\mu}+t}{1+t}$, which is non-decreasing since $\hat{\mu} \in (0, 1]$. Therefore, f reaches its minimal value at the smallest possible value of the argument $t_* = \min\limits_{\boldsymbol{x}\neq\mathbf{0}} \frac{s\boldsymbol{x}^\top \boldsymbol{x}}{\boldsymbol{x}^\top \mathbf{D}\boldsymbol{x}}$. Due to $t_* = \frac{s}{L_{\mathbf{D}}}$, there stands

$$\hat{\mu}_h \geq \frac{\hat{\mu} + s/L_{\mathbf{D}}}{1 + s/L_{\mathbf{D}}}, \tag{53}$$

which, after multiplying both the numerator and the denominator by $L_{\mathbf{D}}/s$ and substituting $s^{-1} = h$, completes the proof. □

Therefore, $\tau_{\mathcal{M},h}$ can be presented as follows

$$\tau_{\mathcal{M},h} = \frac{\hat{\mu} L_{\mathbf{D}} h + 1}{m(L_{\mathbf{D}} h + 1)}, \tag{54}$$

which tends to $\hat{\mu}/m$ as $h \to \infty$ and to $1/m$ as $h \to 0$. Substituting the expression for $\tau_{\mathcal{M},h}$ into (24) and setting the logarithmic derivative of the corresponding criterion equal to zero gives the following equation for the optimal step h_*

$$\frac{L_{\mathbf{D}}}{L_{\mathbf{D}} h_* + 1} - \frac{\hat{\mu} L_{\mathbf{D}}}{\hat{\mu} L_{\mathbf{D}} h_* + 1} + \frac{\mu}{2(h_* \mu + 1)} - \frac{1}{2h_*} = 0. \tag{55}$$

After trivial algebraic transformations, this expression reduces to the quadratic equation

$$\left(L_{\mathbf{D}} \hat{\mu} + 2(\hat{\mu} - 1)\mu\right) h_*^2 + (3\hat{\mu} - 1) h_* + \frac{1}{L_{\mathbf{D}}} = 0. \tag{56}$$

References

Baklanov S, Stefanova M, Lupuleac S (2020) Newton projection method as applied to assembly simulation. Optimization Methods and Software 37(2):577–604, DOI 10.1080/10556788.2020.1818079

Beck A, Teboulle M (2009) A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM Journal on Imaging Sciences 2(1):183–202, DOI 10.1137/080716542

Cheng L, Wang Q, Li J, Ke Y (2015) Variation modeling for fuselage structures in large aircraft digital assembly. Assembly Automation 35(2):172–182, DOI 10.1108/AA-07-2014-069

Devolder O, Glineur F, Nesterov Y (2014) First-order methods of smooth convex optimization with inexact oracle. Math Program 146:37–75, DOI 10.1007/s10107-013-0677-5

Du J, Yue X, Hunt JH, Shi J (2019) Optimal placement of actuators via sparse learning for composite fuselage shape control. Journal of Manufacturing Science and Engineering 141(10):101004, DOI 10.1115/1.4044249

Du J, Cao S, Hunt JH, Huo X, Shi J (2022) A new sparse-learning model for maximum gap reduction of composite fuselage assembly. Technometrics 64(3):409–418, DOI 10.1080/00401706.2022.2050817

d'Aspremont A, Scieur D, Taylor A (2021) Acceleration methods. Foundations and Trends in Optimization 5(1-2):1–245, DOI 10.1561/2400000036

Eliseev A, Lupuleac S, Shinder J, Grigor'ev B (2024a) Simulation of two-way interaction between sealant and structural parts as applied to large-scale aircraft assembly. The International Journal of Advanced Manufacturing Technology DOI 10.1007/s00170-024-13535-6

Eliseev A, Lupuleac S, Shinder J, Grigorev B, Khashba V (2024b) Modeling and analysis of hybrid bonded-bolted joining in large-scale aircraft assembly. The International Journal of Advanced Manufacturing Technology 135:5715–5731, DOI 10.1007/s00170-024-14839-3

Gasnikov A, Gorbunov E, Kovalev D, Mohammed A, Chernousova E (2018) The global rate of convergence for optimal tensor methods in smooth convex optimization. Computer Research and Modeling 10:737–753, DOI 10.20537/2076-7633-2018-10-6-737-753

Gasnikov A, Dvurechensky P, Gorbunov E, Vorontsova E, Selikhanovych D, Uribe CA (2019) Optimal tensor methods in smooth convex and uniformly convex optimization. In: Beygelzimer A, Hsu D (eds) Proceedings of the Thirty-Second Conference on Learning Theory, PMLR, Phoenix, AZ, USA, Proceedings of Machine Learning Research, vol 99, pp 1374–1391, URL https://proceedings.mlr.press/v99/gasnikov19a.html

Gates D (2007) Boeing finds 787 pieces aren't quite a perfect fit. The Seattle Times URL https://www.seattletimes.com/business/boeing-finds-787-pieces-arent-quite-a-perfect-fit/

Giselsson P, Doan D, Keviczky T, De Schutter B, Rantzer A (2013) Accelerated gradient methods and dual decomposition in distributed model predictive control. Automatica 49(3):829–833, DOI 10.1016/j.automatica.2013.01.009

Grippo L, Sciandrone M (2000) On the convergence of the block nonlinear gauss-seidel method under convex constraints. Operations Research Letters 26:127–136, DOI 10.1016/S0167-6377(99)00074-7

Ivanova A, Pasechnyuk D, Grishchenko D, Shulgin E, Gasnikov A, Matyukhin V (2021) Adaptive catalyst for smooth convex optimization. In: Olenev NN, Evtushenko YG, Jaćimović M, Khachay M, Malkova V (eds) Optimization and Applications, Springer International Publishing, Cham, pp 20–37, DOI 10.1007/978-3-030-91059-4_2

Jiang B, Haoyue W, Zhang S (2021) An optimal high-order tensor method for convex optimization. Mathematics of Operations Research 46:1390–1412, DOI 10.1287/moor.2020.1103

Kolossoski O, Monteiro R (2016) An accelerated non-euclidean hybrid proximal extragradient-type algorithm for convex–concave saddle-point problems. Optimization Methods and Software pp 1–29, DOI 10.1080/10556788.2016.1266355

Levitin E, Polyak BT (1966) Constrained minimization methods. USSR Computational Mathematics and Mathematical Physics 6:1–50, DOI 10.1016/0041-5553(66)90114-5

Lin H, Mairal J, Harchaoui Z (2017) Catalyst acceleration for first-order convex optimization: from theory to practice. Journal of Machine Learning Research 18, DOI 10.48550/arXiv.1712.05654

Lin T, Jordan M (2021) A control-theoretic perspective on optimal high-order optimization. Mathematical Programming pp 1–47, DOI 10.1007/s10107-021-01721-3

Lupuleac S, Kovtun M, Rodionova O, Marguet B (2010) Assembly simulation of riveting process. SAE International Journal of Aerospace 2:193 – 198, DOI 10.4271/2009-01-3215

Lupuleac S, Petukhova M, Shinder Y, Bretangol B (2011) Methodology for solving contact problem during riveting process. SAE International Journal of Aerospace 4:952–957, DOI 10.4271/2011-01-2582

Lupuleac S, Shinder Y, Petukhova M, Yakunin S, Smirnov A, Bondarenko D (2013) Development of numerical methods for simulation of airframe assembly process. SAE Int J Aerosp 6(1):101–105, DOI 10.4271/2013-01-2093

Lupuleac S, Shinder J, Churilova M, Zaitseva N, Khashba V, Bonhomme E, Montero-Sanjuan P (2019a) Optimization of automated airframe assembly process on example of a350 s19 splice joint. SAE Technical Paper DOI 10.4271/2019-01-1882

Lupuleac S, Smirnov A, Churilova M, Shinder J, Bonhomme E (2019b) Simulation of body force impact on the assembly process of aircraft parts. In: ASME 2019 International Mechanical Engineering Congress and Exposition, DOI 10.1115/IMECE2019-10635

Lupuleac S, Zaitseva N, Stefanova M, Berezin S, Shinder J, Petukhova M, Bonhomme E (2019c) Simulation of the wing-to-fuselage assembly process. ASME J Manuf Sci Eng 141(6):061009, DOI 10.1115/1.4043365

Lupuleac S, Shinder J, Petukhova M, Zaitseva N, Churilova M (2025) Nonlinear tolerancing: variation simulation for multi-station assembly with compliant parts. The International Journal of Advanced Manufacturing Technology 136:961–981, DOI 10.1007/s00170-024-14884-y

Lutz T, Yue X, Camelio J (2022) Towards a digital twin: Simulation and residual stress analysis in aerospace composite structures assembly. In: Proceedings of the ASME 2022 17th International Manufacturing Science and Engineering Conference (MSEC), ASME, New York, DOI 10.1115/MSEC2022-85439

Marques Alves M (2022) Variants of the a-hpe and large-step a-hpe algorithms for strongly convex problems with applications to accelerated high-order tensor methods. Optimiza-

tion Methods and Software 37:1–31, DOI 10.1080/10556788.2021.2022148
Monteiro RDC, Svaiter BF (2013) An accelerated hybrid proximal extragradient method for convex optimization and its implications to second-order methods. SIAM J Optim 23(2):1092–1125, DOI 10.1137/110833786
Nemirovsky AS, Yudin DB (1983) Problem Complexity and Method Efficiency in Optimization. Wiley, Chichester and New York
Nesterov Y (2012) Efficiency of coordinate descent methods on huge-scale optimization problems. SIAM J Optim 22(2):341–362, DOI 10.1137/100802001
Nesterov Y (2013) Gradient methods for minimizing composite functions. Math Progam 140:125–161, DOI 10.1007/s10107-012-0629-5
Nesterov Y (2018) Lectures on Convex Optimization, 2nd edn. Springer, Cham, DOI 10.1007/978-3-319-91578-4
Nesterov Y, Gasnikov A, Guminov S, Dvurechensky P (2021) Primal–dual accelerated gradient methods with small-dimensional relaxation oracle. Optimization Methods and Software 36(4):773–810, DOI 10.1080/10556788.2020.1731747
Nesterov YE (1983) A method for solving the convex programming problem with convergence rate $o(1/k^2)$. Dokl Akad Nauk SSSR 269(3):543–547
Park J (2022) Fast gradient methods for uniformly convex and weakly smooth problems. Advances in Computational Mathematics 48(34), DOI 10.1007/s10444-022-09943-5
Petukhova M, Lupuleac S, Shinder Y, Smirnov A, Yakunin S, Bretagnol B (2014) Numerical approach for airframe assembly simulation. Journal of Mathematics in Industry 4, DOI 10.1186/2190-5983-4-8
Pogarskaia T, Churilova M, Petukhova M, Petukhov E (2019) Simulation and optimization of aircraft assembly process using supercomputer technologies. Communications in Computer and Information Science 965:367–378, DOI 10.1007/978-3-030-05807-4_31
Qu Z, Zhao C, Wu Y, An L (2026) An efficient equivalent mechanical model for composite panel assembly and its application in assembly process optimization. Composites Part A: Applied Science and Manufacturing 202:109506, DOI 10.1016/j.compositesa.2025.109506
Sadeghi Tabar R, Wärmefjord K, Lindkvist L, Söderberg R (2026) Nonlinear tolerancing through high fidelity variation simulation of deformable parts under frictional contact. CIRP Annals 75(1):155–159, DOI https://doi.org/10.1016/j.cirp.2026.04.084
Song C, Jiang Y, Ma Y (2021) Unified acceleration of high-order algorithms under general hölder continuity. SIAM Journal on Optimization 31:1797–1826, DOI 10.1137/19M1290243
Stefanova M, Yakunin S, Petukhova M, Lupuleac S, Kokkolaras M (2017) An interior-point method-based solver for simulation of aircraft parts riveting. Engineering Optimization 50:1–16, DOI 10.1080/0305215X.2017.1355367
Stefanova M, Minevich O, Baklanov S, Petukhova M, Lupuleac S, Grigor'ev B, Kokkolaras M (2020) Convex optimization techniques in compliant assembly simulation. Optimization and Engineering 21:1665–1690, DOI 10.1007/s11081-020-09493-z
Tseng P (2008) On accelerated proximal gradient methods for convex-concave optimization. SIAM J Optim URL `https://www.mit.edu/~dimitrib/PTseng/papers/apgm.pdf`
Wang Z, Li D, Shen L, Dong X, Zhai Y (2023) Multi-objective optimisation of assembly fixturing layout for large composite fuselage panel reinforced by frames and stringers. The International Journal of Advanced Manufacturing Technology 125(3):1403–1418, DOI 10.1007/s00170-022-10776-1
Wen Y, Yue X, Hunt JH, Shi J (2018) Feasibility analysis of composite fuselage shape control via finite element analysis. Journal of Manufacturing Systems 46:272–281, DOI 10.1016/j.jmsy.2018.01.008
Wen Y, Yue X, Hunt JH, Shi J (2019) Virtual assembly and residual stress analysis for the composite fuselage assembly process. Journal of Manufacturing Systems 52:55–62, DOI 10.1016/j.jmsy.2019.04.001
Xue B, Liu L, Xiao H, Liang H, Chang M (2026) Analysis of orbital joints in composite airplane fuselage with omega stringer. In: Yang RC (ed) Proceedings of the 7th International Conference on Innovative and Smart Materials, Springer Nature Singapore, Singapore, pp 97–108, DOI 10.1007/978-981-95-6171-1_9
Zaitseva N, Shinder J, Lupuleac S, Khashba V, Titova M, Baklanov S (2025) Decomposition strategies for the contact problem in fuselage assembly process modeling.

In: Optimization and Applications, Springer International Publishing, Cham, DOI 10.1007/978-3-032-13127-0_15