Journal of Global Optimization **14:** 117–135, 1999.

Comparing Nonsmooth Nonconvex Bundle Methods in Solving Hemivariational Inequalities

M.M. MÄKELÄ[1], M. MIETTINEN[1], L. LUKŠAN[2] and J. VLČEK[2]
[1]*Laboratory of Scientific Computing, Department of Mathematics, University of Jyväskylä, P.O. Box 35, FIN-40351 Jyväskylä, Finland (email: mmiettin@math.jyu.fi);* [2]*Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 2, 182 07 Prague 8, Czech Republic*

(Received 16 October 1997; accepted in revised form 12 January 1998)

Abstract. Hemivariational inequalities can be considered as a generalization of variational inequalities. Their origin is in nonsmooth mechanics of solid, especially in nonmonotone contact problems. The solution of a hemivariational inequality proves to be a substationary point of some functional, and thus can be found by the nonsmooth and nonconvex optimization methods. We consider two type of bundle methods in order to solve hemivariational inequalities numerically: proximal bundle and bundle-Newton methods. Proximal bundle method is based on first order polyhedral approximation of the locally Lipschitz continuous objective function. To obtain better convergence rate bundle-Newton method contains also some second order information of the objective function in the form of approximate Hessian. Since the optimization problem arising in the hemivariational inequalities has a dominated quadratic part the second order method should be a good choice. The main question in the functioning of the methods is how remarkable is the advantage of the possible better convergence rate of bundle-Newton method when compared to the increased calculation demand.

Key words: Bundle methods, Hemivariational inequalities, Nondifferentiable programming, Nonmonotone contact problems, Substationary points

1. Introduction

Hemivariational inequalities introduced by Panagiotopoulos are generalizations of variational inequalities. By means of them, problems involving nonmonotone and multivalued constitutive laws and boundary conditions can be defined mathematically. In many cases hemivariational inequalities can be reformulated as substationary point problems of the corresponding nonsmooth nonconvex energy functionals. For mathematical theory and the applications of hemivariational inequalities we refer to [22, 23].

The aim of this paper is to apply nonsmooth nonconvex optimization methods for the numerical solution of hemivariational inequalities and analyse their efficiency. As a typical example of hemivariational inequalities we consider a laminated composite structure under loading when the binding material between laminae obeys a nonmonotone multivalued law. This kind of a mechanical problem has been investigated numerically in [21, 26]. There it has been replaced by a sequence of convex subproblems which have been solved by using convex minimization methods.

The discretization of the considered problem is realized by the finite element method scheme for nonmonotone multivalued differential inclusions presented in [17–19]. This scheme has been proved to be mathematically well-posed: stable and convergent.

Bundle methods are at the moment the most promising methods for nonsmooth optimization. Their origin is the classical cutting plane method of [4] and [7] and they are based on the piecewise linear approximation of the objective function. Due to the numerical experiments the proximal bundle methods (see [10, 25, 16]) seem to work in the most efficient and reliable way. They can be called also to diagonal variable metric methods, since a stabilizing quadratic term in form of diagonal matrix was added to the polyhedral approximation in order to accumulate some second order information about the curvature of the objective function.

However, the development of 'real' second order method has been fascinating the researchers of nonsmooth optimization during its whole history. Several attempts have been done in order to exploit the second order subderivative information. Already in his pioneering work [11] Lemaréchal derived a version of variable metric bundle method utilizing the classical BFGS secant updating formula from smooth optimization. Due to the disappointing numerical results in [12] this idea was buried nearly for two decades. Several modifications of the variable metric concepts have been proposed for example in [3, 6, 13, 20]. According to very limited numerical experiments (see for example [6]) it seems that the variable metric bundle methods works fairly well. However, when proportion the results to the extra computational efforts needed with the full matrix algebra they do not offer substantial advancement in numerical solution process.

More recently a new second order approach has been proposed in [15], where the bundle-Newton method was introduced. The main difference compared to the earlier methods was the inclusion of second order information directly to the model, in other words the piecewise linear model was replaced by piecewise quadratic model. In numerical tests of [15] it turns out to be very effective for quadratic type of problems.

The aim of this paper is to compare the proximal bundle method and the bundle-Newton method in solving hemivariational inequalities. Since the optimization problem arising in the solution of hemivariational inequalities has a dominated quadratic part the second order method should be a good choice. The main question in the functioning of the method is how remarkable is the advantage of the possible better convergence rate of the bundle-Newton method when compared to the increased calculation demand. The numerical results indicate the superiority of the bundle-Newton method for the problems having a quadratic nature.

The paper is organized as follows. In Section 2 we introduce the considered optimization problem and give some definitions needed in continuation. The Sections

3 and 4 are devoted to the proximal bundle method and bundle-Newton method, respectively. In Section 5 the test hemivariational inequality problem is introduced and finally in Section 6 we give some numerical results.

2. Nonsmooth and nonconvex optimization problem

We consider the following nonsmooth and nonconvex optimization problem

$$\begin{cases} \text{minimize} & f(x) \\ \text{subject to} & x \in K, \end{cases} \tag{P}$$

where the objective function f from R^n to R is supposed to be locally Lipschitz continuous function. The feasible set K has a more specific structure, i.e.

$$K = \{x \in R^n \mid x^l \leqslant x \leqslant x^u\},$$

where x^l and x^u are the lower and upper bounds for variables, respectively. We suppose that at each $x \in K$ we can evaluate the function value $f(x)$ and an arbitrary subgradient $g(x)$ from the subdifferential of Clarke (see [5])

$$\partial f(x) = \text{conv}\{\lim_{i\to\infty} \nabla f(x^i) \mid x^i \to x \text{ and } \nabla f(x^i) \text{ exists}\}.$$

The *normal cone* of K at $x \in K$ can be defined (see [16]) by

$$N_K(x) = \text{cl}\left\{\bigcup_{\lambda\geqslant 0} \lambda\, \partial d_K(x)\right\},$$

where $d_K : R^n \to R$ is the *distance function* of K, i.e.

$$d_K(x) = \inf\left\{\|x - c\| \mid c \in K\right\}.$$

DEFINITION 1. *Suppose that $f : R^n \to R$ is locally Lipschitz continuous. Then x^* is called a substationary point of the problem (P) if*

$$0 \in \partial f(x^*) + N_K(x^*). \tag{1}$$

Now we can formulate the following necessary optimality condition.

THEOREM 1. *Every local minimizer of the problem (P) is substationary.*

For the proof we refer to [16].

In the following sections we describe two methods: proximal bundle method and bundle-Newton method for finding local minimizers (and thus substationary) points of the problem (P). For the convergence of the methods we need the following semismoothness assumption due to [2].

DEFINITION 2. *The function* $f : R^n \to R$ *is said to be upper semismooth, if for any* $x \in R^n$, $d \in R^n$ *and sequences* $g_i \subset R^n$ *and* $t_i \subset (0, \infty)$ *satisfying* $g_i \in \partial f(x + t_i d)$ *and* $t_i \downarrow 0$, *one has*

$$\limsup_{i\to\infty} g_i^T d \geqslant \liminf_{i\to\infty} [f(x + t_i d) - f(x)]/t_i. \tag{2}$$

3. Proximal bundle method

In this section we shortly describe the ideas of the proximal bundle method for nonsmooth and nonconvex minimization. For more details we refer to [10, 25, 16].

3.1. DIRECTION FINDING

Our aim is to produce a sequence $\{x_k\}_{k=1}^{\infty} \subset R^n$ converging to some local minimum of the problem (P). Suppose that the starting point x_1 is feasible and at the kth iteration of the algorithm we have the current iteration point x_k and some trial points $y_j \in R^n$ (from past iterations) and subgradients $g_j \in \partial f(y_j)$ for $j \in J_k$, where the index set J_k is a nonempty subset of $\{1, \ldots, k\}$.

The idea behind the proximal bundle method is to approximate the objective function below by a piecewise linear function, in other words, we replace f by so called *cutting-plane model*

$$\hat{f}_k(x) := \max_{j\in J_k} \{f(y_j) + g_j^T (x - y_j)\}, \tag{3}$$

which equivalently can be written in the form

$$\hat{f}_k(x) = \max_{j\in J_k} \{f(x_k) + g_j^T (x - x_k) - \alpha_j^k\}, \tag{4}$$

with the *linearization error*

$$\alpha_j^k := f(x_k) - f(y_j) - g_j^T (x_k - y_j) \quad \text{for all } j \in J_k. \tag{5}$$

Note, that in convex case the subdifferential can be rewritten as (see [16])

$$\partial f(x) = \{g(x) \in R^n \mid f(y) \geqslant f(x) + g(x)^T (y - x) \text{ for all } y \in R^n\}.$$

Then it is easy to prove that

$$\hat{f}_k(x) \leqslant f(x) \quad \text{for all } x \in R^n \quad \text{and} \quad \alpha_j^k \geqslant 0 \quad \text{for all } j \in J_k. \tag{6}$$

In other words, if f is convex, then the cutting-plane model $\hat{f}_k$ is an under estimate for f and the nonnegative linearization error α_j^k measures how good an approximation the model is to the original problem. In nonconvex case these facts are not valid anymore: α_j^k may have a tiny (or even negative) value, although the trial point y_j lies far away from the current iteration point x_k and thus the corresponding subgradient g_j is useless. For these reasons the linearization error α_j^k is replaced by so called *subgradient locality measure* (cf. [8])

$$\beta_j^k := \max\{|\alpha_j^k|, \gamma(s_j^k)^2\}, \tag{7}$$

where $\gamma \geqslant 0$ is the *distance measure parameter* ($\gamma = 0$ if f is convex) and

$$s_j^k := \|x_j - y_j\| + \sum_{i=j}^{k-1} \|x_{i+1} - x_i\| \tag{8}$$

is the *distance measure* estimating

$$\|x_k - y_j\| \tag{9}$$

without the need to store the trial points y_j. Then obviously $\beta_j^k \geqslant 0$ for all $j \in J_k$ and $\min_{x\in K} \hat{f}_k(x) \leqslant f(x_k)$, since

$$\min_{x\in K} \hat{f}_k(x) \leqslant \hat{f}_k(x_k) = f(x_k) - \max_{j\in J_k} \beta_j^k \leqslant f(x_k). \tag{10}$$

In order to calculate the search direction $d_k \in R^n$ we replace the original problem (P) by the cutting plane model

$$\begin{cases} \text{minimize} & \hat{f}_k(x_k + d) + \frac{1}{2}u_k d^T d \\ \text{subject to} & x_k + d \in K, \end{cases} \tag{CP}$$

where the regularizing quadratic penalty term $1/2u_k d^T d$ is added to guarantee the existence of the solution d_k and keep the approximation local enough. The weighting parameter $u_k > 0$ was added to improve the convergence rate and to accumulate some second order information about the curvature of f around x_k. It was adapted from the proximal point algorithm by [24] and [1] and was first time used in [10] and [25].

Notice, that the problem (CP) still is a nonsmooth optimization problem. However, due to piecewise linear nature it can be rewritten as a (smooth) quadratic programming subproblem finding the solution $(d_k, v_k) \in R^{n+1}$ of

$$\begin{cases} \text{minimize} & v + \frac{1}{2}u_k d^T d \\ \text{subject to} & -\beta_j^k + g_j^T d \leqslant v \quad \text{for all } j \in J_k \quad \text{and} \quad x_k + d \in K. \end{cases} \tag{QP}$$

3.2. LINE SEARCH

In the previous section we calculated the search direction d_k. Next we consider the problem of determining the step size into that direction. We assume that $m_L \in$

$(0, 1/2)$, $m_R \in (m_L, 1)$ and $\bar{t} \in (0, 1]$ are fixed line search parameters. First we shall search for the largest number $t_L^k \in [0, 1]$ such that $t_L^k \geqslant \bar{t}$ and

$$f(x_k + t_L^k d_k) \leqslant f(x_k) + m_L t_L^k v_k, \tag{11}$$

where v_k is the predicted amount of descent and it holds

$$v_k = \hat{f}_k(x_k + d_k) - f(x_k) < 0$$

due to (11). If such a parameter exists we take a *long serious step*

$$x_{k+1} := x_k + t_L^k d_k \quad \text{and} \quad y_{k+1} := x_{k+1}.$$

Otherwise, if (11) holds but $0 < t_L^k < \bar{t}$ then a *short serious step*

$$x_{k+1} := x_k + t_L^k d_k \quad \text{and} \quad y_{k+1} := x_k + t_R^k d_k$$

is taken and if $t_L^k = 0$ we take a *null step*

$$x_{k+1} := x_k \quad \text{and} \quad y_{k+1} := x_k + t_R^k d_k,$$

where $t_R^k > t_L^k$ is such that

$$-\beta_{k+1}^{k+1} + g_{k+1}^T d_k \geqslant m_R v_k. \tag{12}$$

In long serious step there occurs a significant decrease in the value of the objective function. Thus there is no need for detecting discontinuities in the gradient of f, and so we set $g_{k+1} \in \partial f(x_{k+1})$. In short serious steps and null steps there exists discontinuity in the gradient of f. Then the requirement (12) ensures that x_k and y_{k+1} lie on the opposite sides of this discontinuity and the new subgradient $g_{k+1} \in \partial f(y_{k+1})$ will force a remarkable modification of the next search direction finding problem. In what follows we are using the line search algorithm presented in [16]. The convergence proof of the algorithm is assumed f to be upper semismooth (see (2)).

The iteration is terminated if

$$v_k \geqslant -\varepsilon_s, \tag{13}$$

where $\varepsilon_s > 0$ is a final accuracy tolerance supplied by the user.

3.3. WEIGHT UPDATING

One of the most important questions concerning proximal bundle method is the choice of the weight u_k. The simplest strategy might be to keep it constant $u_k \equiv u_{fix}$. This, however, leads to several difficulties. Due to Theorem 4 we observe the following:

If u_{fix} is very large, we shall have small $|v_k|$ and $\|d_k\|$, almost all steps are serious and we have slow descent.
If u_{fix} is very small, we shall have large $|v_k|$ and $\|d_k\|$, and each serious step will be followed by many null steps.

Therefore, we keep the weight as a variable and update it when necessary. For updating u_k we use the safeguarded quadratic interpolation algorithm due to [10].

4. Bundle-Newton method

Next we describe the main ideas of the second order bundle-Newton method. For more details we refer to [15].

4.1. DIRECTION FINDING

We suppose that at each $x \in K$ we can evaluate, in addition to the function value $f(x)$ and an arbitrary subgradient $g(x) \in \partial f(x)$, also an $n \times n$ symmetric matrix $G(x)$ approximating the Hessian matrix $\nabla^2 f(x)$. For example, at the kink point y of piecewise twice differentiable function we can take $G(y) = \nabla^2 f(x)$, where x is 'infinitely close' to y.

Instead of piecewise linear cutting-pane model (3) we introduce a piecewise quadratic model of the form

$$\tilde{f}_k(x) := \max_{j \in J_k} \left\{ f(y_j) + g_j^T(x - y_j) + \frac{1}{2}\varrho_j(x - y_j)^T G_j(x - y_j) \right\}, \tag{14}$$

where $G_j = G(y_j)$ and $\varrho_j \in [0, 1]$ is some damping parameter. The model (14) can again equivalently be written as

$$\tilde{f}_k(x) = \max_{j \in J_k} \left\{ f(x_k) + g_j^T(x - x_k) + \frac{1}{2}\varrho_j(x - x_k)^T G_j(x - x_k) - \alpha_j^k \right\} \tag{15}$$

and for all $j \in J_k$ the linearization error takes now the form

$$\alpha_j^k := f(x_k) - f(y_j) - g_j^T(x_k - y_j) - \frac{1}{2}\varrho_j(x_k - y_j)^T G_j(x_k - y_j). \tag{16}$$

Note that now even in the convex case α_j^k might be negative. Therefore we replace the linearization error (16) again by the subgradient locality measure (7) and we remain the property (see [15])

$$\min_{x \in K} \tilde{f}_k(x) \leqslant f(x_k). \tag{17}$$

The search direction $d_k \in R^n$ is now calculated as the solution of

$$\begin{cases} \text{minimize} & \tilde{f}_k(x_k + d) \\ \text{subject to} & x_k + d \in K. \end{cases} \tag{CN}$$

Note, that since the model already has second order information no regularizing quadratic terms are needed like in (CP). The problem (CN) is transformed to a nonlinear programming problem, which is then solved by a recursive quadratic programming method (see [15]). If we denote

$$g_j^k := g_j + \varrho_j G_j(x_k - y_j),$$

this procedure leads to a quadratic programming subproblem finding the solution $(d_k, v_k) \in R^{n+1}$ of

$$\begin{cases} \text{minimize} & v + \frac{1}{2} d^T W_k d \\ \text{subject to} & -\beta_j^k + (g_j^k)^T d \leqslant v \quad \text{for all } j \in J_k \quad \text{and} \quad x_k + d \in K, \end{cases} \tag{QN}$$

where

$$W_k := \sum_{j \in J_{k-1}} \lambda_j^{k-1} \varrho_j G_j$$

and λ_j^{k-1} for $j \in J_{k-1}$ are the Lagrange multipliers of (QN) from the previous iteration $k - 1$. In calculations W_k is replaced by its positive definite modification, if necessary.

4.2. LINE SEARCH

The line search operation of the bundle-Newton method is following the same principles than in Section 3.2 for proximal bundle method. The only remarkable difference occurs in the termination condition for short and null steps, in other words (12) is replaced by two conditions

$$-\beta_{k+1}^{k+1} + (g_{k+1}^{k+1})^T d_k \geqslant m_R v_k \tag{18}$$

and

$$\|x_{k+1} - y_{k+1}\| \leqslant C_S, \tag{19}$$

where $C_S > 0$ is a parameter supplied by the user.

The bundle-Newton method is using the line search algorithm presented in [15]. The convergence proof of the algorithm is assumed f to be upper semismooth (see (2)).

5. Formulation of the problem

5.1. CONTINUOUS PROBLEM

We consider a two-dimensional laminated composite structure consisting of two elastic laminae and an interlayer binding material under loading (see Figure 1).

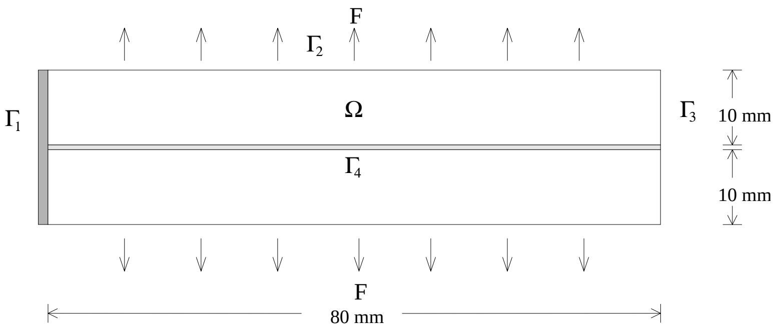

Figure 1. Laminated composite structure under loading.

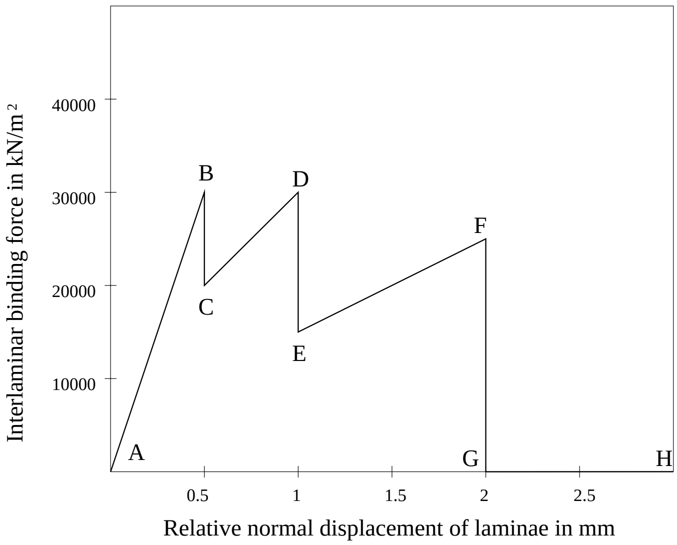

Figure 2. Nonmonotone adhesive force between laminae.

The mechanical behaviour of the binding material, the interlaminar bonding forces versus the corresponding relative displacements of the laminae, is depicted by Figure 2. This relation is typically nonmonotone and multivalued. Therefore, this kind of mechanical problems lead to the hemivariational inequalities or in the energy formulation to the substationary point problems of nonconvex, nonsmooth functionals. The similar mechanical structures have been investigated in [21, 26].

Due to the symmetry of the structure and by assuming that the forces applied to the upper and lower part of the structure are the same it is enough to study the upper lamina. Next, we formulate mathematically the problem. We denote by $\Omega \subset R^2$ the upper lamina in its undeformed state. The Lipschitz boundary Γ of Ω consists of four nonoverlapping open subsets Γ_1, Γ_2, Γ_3 and Γ_4. We denote by $\sigma = (\sigma_{ij})_{i,j=1}^2$ the stress tensor, $\varepsilon = (\varepsilon_{ij})_{i,j=1}^2$ the strain tensor, $u = (u_i)_{i=1}^2$ the displacement,

$n = (n_i)_{i=1}^2$ the outward unit normal vector to Γ and $S = (S_i = \sigma_{ij} n_j)_{i=1}^2$ the boundary force.

Assuming that the deformations are small the lamina obeys the Hooke's law of form

$$\sigma_{ij} = C_{ijkl}\varepsilon_{kl}(u), \quad \text{where} \quad \varepsilon_{kl}(u) = \frac{1}{2}\left(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k}\right) \tag{20}$$

and $C = (C_{ijkl})_{i,j,k,l=1}^2$ is the elasticity tensor satisfying the usual symmetry and ellipticity conditions. The equation of the equilibrium state of Ω is as follows:

$$\sigma_{ij,j} = 0 \quad \text{in } \Omega, \quad i = 1, 2, \tag{21}$$

since there are no volume forces.

Next, we define the boundary conditions. On Γ_1 we have given displacements

$$u(x) = 0 \quad \text{on } \Gamma_1 \tag{22}$$

and on both Γ_2 and Γ_3 given boundary forces

$$S(x) = (0, F) \quad \text{on } \Gamma_2, \tag{23}$$

$$S(x) = 0 \quad \text{on } \Gamma_3, \tag{24}$$

where F is a constant force.

Furthermore, on Γ_4 we have in the tangential direction a given boundary force

$$S_1(x) = 0 \qquad \text{on } \Gamma_4. \tag{25}$$

On the other hand, due to the binding interlayer material it holds a nonmonotone, multivalued boundary condition (see Figure 2) expressed by means of the subdifferential ∂j of a locally Lipschitz function j

$$-S_2(x) \in \partial j(u_2(x)) \quad \text{on } \Gamma_4, \tag{26}$$

and because of the nonpenetration of the laminae a unilateral boundary condition

$$u_2 \geqslant 0 \quad \text{on } \Gamma_4 \tag{27}$$

in the normal direction. Note that we have to scale by $1/2$ the x_1-axis of the graph in Figure 2 in order to get the nonmonotone law of (26), because in Figure 2 we have relative displacements of the laminae.

By employing Green–Gauss theorem and taking into account both the equilibrium equation (21) and the boundary conditions (22)–(27) we obtain the following hemivariational inequality: Find $u \in K$ such that

$$a(u, v - u) + \int_{\Gamma_4} j^\circ(u_2; v_2 - u_2)d\Gamma \geqslant \int_{\Gamma_2} F(v_2 - u_2)d\Gamma \quad \forall v \in K, \tag{28}$$

where K is a set of kinematical admissible displacements defined by

$$K = \{v \in X : v_2 \geqslant 0 \text{ on } \Gamma_4\}, \tag{29}$$

$$X = \{v \in (H^1(\Omega))^2 : v = 0 \text{ on } \Gamma_1\} \tag{30}$$

($H^1(\Omega)$ denotes the Sobolev space), j° is the generalized directional derivative of j and a the bilinear form of the linear elasticity defined by

$$a(u, v) = \int_\Omega C_{ijhk}\varepsilon_{ij}(u)\varepsilon_{hk}(v)d\Omega. \tag{31}$$

Due to the symmetry of the bilinear form a the potential energy of the considered mechanical system has the form:

$$f(u) = \frac{1}{2}a(u, u) + \int_{\Gamma_4} j(u_2)d\Gamma - \int_{\Gamma_2} Fu_2 d\Gamma. \tag{32}$$

Then, we can formulate the following problem: Find $u \in K$ such that u is a substationary point (cf. (1)) of f on K, i.e.

$$0 \in \partial f(u) + N_K(u), \tag{33}$$

where $N_K(u)$ is the normal cone to the nonempty, closed and convex set K at u. It can be shown that the problem (28) has at least one solution (see [22]), and that the solutions of the inclusion (33) are also the solutions of the inequality (28) (see [19]).

5.2. DISCRETE PROBLEM

We apply the finite element approximation scheme developed in [17, 18]. Let h be a discretization parameter related to the mesh size of the triangulation $\mathcal{T}_h$ of Ω (see Figure 3). The set K is approximated by the set of piecewise linear functions over the triangulation $\mathcal{T}_h$ defined by

$$K_h = \{v \in X_h : v_2 \geqslant 0 \text{ on } \Gamma_4\}, \tag{34}$$

$$X_h = \{v \in (C(\overline{\Omega}))^2 : v|_T \in (P_1(T))^2\ \forall T \in \mathcal{T}_h, v = 0 \text{ on } \Gamma_1\}. \tag{35}$$

For the approximation of the bilinear form a and the boundary integral $\int_{\Gamma_2} F(v_2 - u_2)d\Gamma$ we use appropriate numerical integration formulae (this is standard in the finite element method). The nonmonotone term in (28) is approximated by the following numerical integration formula

$$\int_{\Gamma_4} j^\circ(u_2; v_2 - u_2)d\Gamma \approx \sum_{i\in I} c_i j^\circ(u_2(x_i); v_2(x_i) - u_2(x_i)), \tag{36}$$

where $\{x_i\}_{i\in I}$ is the set of the nodal points of the triangulation $\mathcal{T}_h$ on Γ_4 and $\{c_i\}_{i\in I}$ are the coefficients of the integration formula.

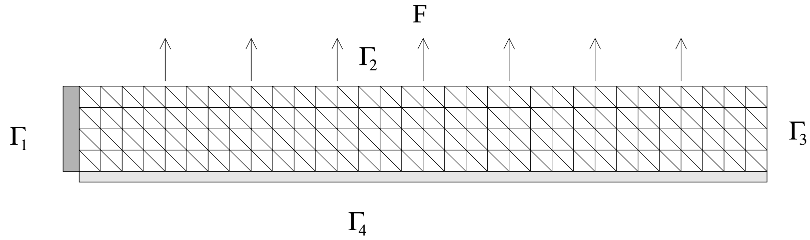

Figure 3. 32×4 triangulation $\mathcal{T}_h$ of Ω.

Let us make the identification X_h with the corresponding subset of R^{2n}, i.e. the displacement function v is identified with the displacement vector of the nodal points (n is the number of the nodal points). Then the discrete problem in the matrix form is as follows: Find $u \in K_h$ such that

$$u^T A(v-u) + \sum_{i\in I} c_i j^\circ(u_i; v_i - u_i) \geqslant F^T(v-u) \quad \forall v \in K_h, \tag{37}$$

where A is the stiffness matrix of the structure, F the (discrete) load vector and I the set of the indices corresponding to the x_2-displacements of the nodal points on Γ_4.

Then we can formulate the discrete substationary point problem on K_h: Find $u \in K_h$ such that

$$0 \in \partial f_h(u) + N_{K_h}(u), \tag{38}$$

where

$$f_h(u) = \frac{1}{2}u^T Au + \sum_{i\in I} c_i j(u_i) - F^T u. \tag{39}$$

Because of Corollary 1 of Proposition 2.3.3 in [5] and the terms $j(u_i)$, $i \in I$, are independent of each other we have that

$$\partial f_h(u) = Au + \sum_{i\in I} c_i \partial j(u_i) - F. \tag{40}$$

Now it holds that every solution of (38) is a solution of the problem (37) (see [19]). Also due to the results in [18, 19] we know that the discrete problem (38) is solvable and its solutions converge in subsequences to the solutions of the continuous problem (33).

REMARK . The functional f_h is upper semismooth. Indeed, let $u, d \in R^{2n}$ be given. From Definition 2 of the upper semismoothness we see that it is enough to study the restriction of f_h onto the line $L = \{u + td : t \in R\}$, on which f_h is smooth except finite number of points (cf. (39) and Figure 2). If $f_h|_L$ is smooth at

u, it is smooth also on some small (one-dimensional) neighborhood of u and the condition

$$\limsup_{i\to\infty} g_i^T d \geqslant \liminf_{i\to\infty} [f(x + t_i d) - f(x)]/t_i \tag{41}$$

is trivially satisfied. On the other hand, if $f_h|_L$ is nonsmooth in u, there exists small neighborhood of u in which u is the only nonsmooth point. Therefore, $f_h|_L$ is continuously differentiable in this set except at the point u and the classical one-sided derivatives exist at u implying (41).

Since the nonlinear behaviour and the constraints of the problem have an effect only on the x_2-displacements on the nodes on Γ_4, we can make our computation much more effective by applying the method of condensation of unknowns. Assuming that m components corresponding to the index set I are listed first, we have the following decomposition of the matrix A and the vectors F and u

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \qquad F = \begin{pmatrix} F^1 \\ F^2 \end{pmatrix} \qquad u = \begin{pmatrix} u^1 \\ u^2 \end{pmatrix},$$

where A_{11} is an $m \times m$ matrix and $F^1, u^1 \in R^m$. Then the elimination of $u^2 (\in R^{2n-m})$ from (37) leads to the following discrete potential functional

$$\tilde{f}_h(u^1) = \frac{1}{2}(u^1)^T \tilde{A} u^1 + \sum_{i=1}^{m} c_i j((u^1)_i) - (\tilde{F})^T u^1, \tag{42}$$

where $\tilde{A} = A_{11} - A_{12}(A_{22})^{-1} A_{21}$ called the Schur complement and $\tilde{F} = F^1 - A_{12}(A_{22})^{-1} F^2$. Hence, the eliminated substationary point problem is formulated as follows: Find $u^1 \in K_h$ such that

$$0 \in \tilde{A} u^1 + \sum_{i=1}^{m} c_i \partial j((u^1)_i) - \tilde{F} + N_{K_h}(u^1), \tag{43}$$

where K_h is interpreted as a set of R^m.

6. Numerical results

The optimization algorithms has been implemented in Fortran 77 and the test runs have been performed on an HP9000/J280 (180 MHz) computer. The tested optimization codes are presented in Table 1.

All the codes are utilizing the subgradient aggregation strategy of [8] to keep the storage requirements bounded. The code PB is employing the quadratic solver QPDF4, which is based on the dual active set method derived in [9], while the codes PBL and BNL are using the solver ULQDF1 implementing the dual projected gradient method proposed in [14].

Table 1. Tested codes and software

Code	Software Package	Author(s)	Method
PB	NSOLIB	Mäkelä	Proximal bundle
PBL	UFO	Lukšan &Vlček	Proximal bundle
BNL	UFO	Lukšan &Vlček	Bundle-Newton

Table 2. Eliminated with discretization 16×2

	Load=26 200 kN/m^2				Load=27 000 kN/m^2			
Code	Ni	Nf	CPU	f	Ni	Nf	CPU	f
PB	1942	2209	10.76	−0.830188	1698	8614	15.74	−8.127245
PBL	207	210	1.13	−0.830189	1421	1443	3.80	−8.127215
BNL	4	5	0.69	−0.830189	23	24	0.81	−8.127245

The above optimization codes have been applied to the mechanical structure of Figure 1 under eleven constant loadings (F=20 000, 21 000, 22 000, 23 000, 24 000, 25 000, 26 200, 27 000, 28 000, 29 000, 30 000 kN/m^2). In calculation we have used the incremental procedure: the loading of the structure is increased uniformly and as an initial guess for the solution of the next load it is used the solution of the previous load. The bigger increment from 25 000 kN/m^2 to 26 200 kN/m^2 is due to the fact that the structure is very sensitive to the increase of the loading between 25 000-27 000 kN/m^2 where the partial (branches B–G in Figure 2) and the complete delamination (branch G–H in Figure 2) take place. With the load 26 200 kN/m^2 we can illustrate the case in which only the partial delamination occurs. We have applied the plane stress model with the elasticity modulus $E = 1.378 \cdot 10^8$ kN/m^2, the Poisson's ratio $\nu = 0.3$ and the thickness $t = 5$ mm.

The calculated results with different discretizations are presented in Tables 2–6, in which Ni denotes the number of iterations, Nf denotes the number of objective

Table 3. Eliminated with discretization 32×4

	Load=26 200 kN/m^2				Load=27 000 kN/m^2			
Code	Ni	Nf	CPU	f	Ni	Nf	CPU	f
PB	2784	8627	48.45	−13.96062*	1248	4715	23.34	−15.21955
PBL	426	434	3.15	−0.871201	3007	3063	14.16	−15.21929
BNL	13	14	1.61	−0.871203	16	17	1.70	−15.21955

Table 4. Eliminated with discretization 64×8

Code	Load=26 200 kN/m^2				Load=27 000 kN/m^2			
	Ni	Nf	CPU	f	Ni	Nf	CPU	f
PB	4053	13480	198.33	−17.00291*	1909	7411	107.00	−18.45613
PBL	799	810	31.23	−0.884754	6131	6255	85.33	−18.45483
BNL	15	16	24.71	−0.884754	14	15	24.67	−18.45613

Table 5. Eliminated with discretization 128×16

Code	Load=26 200 kN/m^2				Load=27 000 kN/m^2			
	Ni	Nf	CPU	f	Ni	Nf	CPU	f
PB	6414	23681	1372.03	−17.93623*	3073	12287	939.48	−19.45147
PBL	184	188	551.16	−0.809380**	11653	11875	873.10	−19.44819**
BNL	17	19	552.69	−0.888642	14	15	557.94	−19.45147

function (and also subgradient) evaluations, CPU the used total computer time in seconds and f is the objective function value at termination. Note that the elimination, the computation of the Schur complement $\tilde{A}$ and the new load vector $\tilde{F}$, for the discretization 16×2, 32×4, 64×8 and 128×16 is taking about 0.67, 1.35, 23.1 and 545.0 seconds, respectively. The stopping criterions are chosen to be comparable, in other words the relative accuracy in the optimal objective function value is about 6 digits.

From the numerical results in Tables 2–6 we can conclude the superiority of the bundle-Newton code BNL. In all cases it used less computing resources and found the local minimum in the most reliable way. Due to the matrix operations the individual iteration is more costly than in proximal bundle methods. However, the total amount of iterations stays very low, and thus the used CPU time of BNL is always less than with the other codes.

Table 6. Not eliminated with discretization 32×4

Code	Load=26 200 kN/m^2				Load=27 000 kN/m^2			
	Ni	Nf	CPU	f	Ni	Nf	CPU	f
PB	4458	4474	156.97	−0.880013	40112	40128	1458.51	−15.22886
PBL	2957	3004	40.96	−0.880011	22606	22987	312.08	−15.21907**
BNL	10	12	12.27	−0.880022	14	17	17.79	−15.22891

Table 7. Average error of the normal displacements on Γ_4 by BNL

Discretization	Load=26 200 kN/m^2 (Error in mm)	Load=27 000 kN/m^2 (Error in mm)
16×2	0.027	2.7
32×4	0.0094	1.0
64×8	0.0026	0.24
128×16	$*$	$*$

Note, that the iteration number and the function evaluations of BNL do not depend on the dimension of the problem. On the other hand, when the size of problem is doubled, BPL needed about two times and PB about 1.6 times more iterations and function evaluations.

When comparing the CPU times versus discretization the behaviour is nonlinear and the multiplier is growing exponentially. We have to remember, that especially in the large problems the elimination is taking most of the time and the differences do not seem to be so remarkable. For example for the discretization 128×16 the optimization time of BNL is only about 2% of the total time.

By comparing the Tables 3 and 6 we can see the influence of the elimination; the solution of the problem without elimination is taking nearly ten times more CPU time than the eliminated one. Note that especially for the proximal bundle methods the elimination is essential.

Although the codes PB and PBL are realizations of the same method, the difference in their functioning is remarkable. PBL is clearly more effective when comparing the function evaluations and CPU times. However, it has slight difficulties to reach the desired accuracy in larger problems (see the function values denoted by $**$ in Tables 5–6). Notice also, that PB converges to the different local minimum with the load 26 200 kN/m^2 (see the function values denoted by $*$ in Tables 3–5). When we changed the starting point to be zero in all the components it found the same optimum as the other codes. This behaviour is due the fact that the delamination takes place between the loads 25 000-27 000 kN/m^2 making the structure very unstable and sensitive to the increase of the load.

In Table 7 we have listed the average errors of the normal displacements on Γ_4 with the loads 26 200 kN/m^2 and 27 000 kN/m^2 obtained by BNL. As an exact solution it is used the solution obtained by the discretization 128×16. Table 7 indicates that the applied approximation scheme has a very good convergence rate and the results calculated by BNL are reliable.

In Figures 4 and 5 it is presented the normal displacements of the upper lamina (the relative displacements of laminae are twice larger) and the binding forces between laminae. Further, Figure 6 illustrates the progress of delamination: With the load 23 000 kN/m^2 there is no damage of the binding material, all the nodes

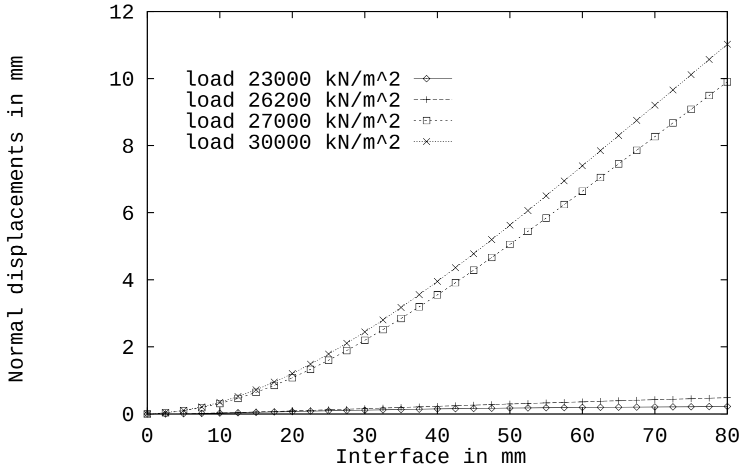

Figure 4. Normal displacements of the interface of the upper lamina.

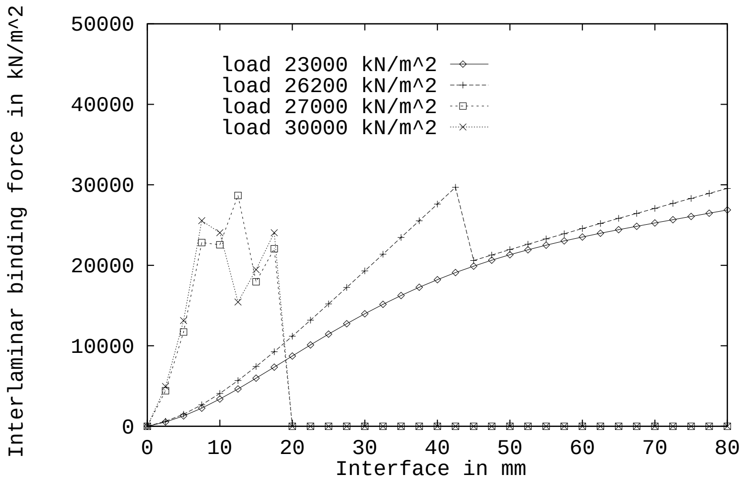

Figure 5. Nonmonotone adhesive force between laminae.

are on the branch A-B. The partial delamination occurs when the load is increased to 26 200 kN/m^2, some of the nodes are on the branch C-D, and the complete delamination when the load is greater than 27 000 kN/m^2, most of the nodes are on the branch G-H.

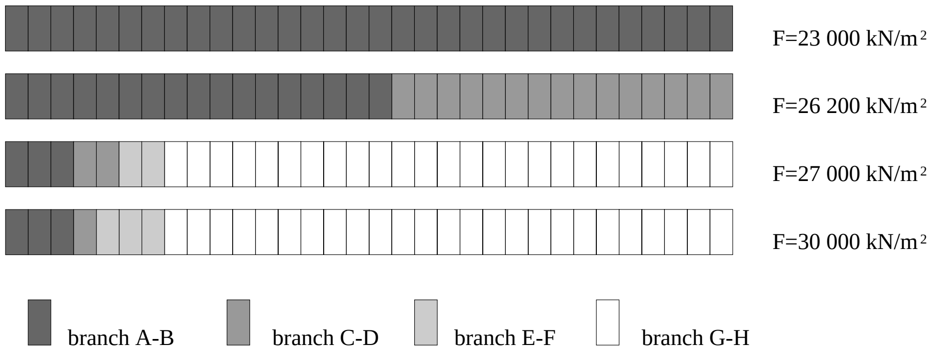

Figure 6. Progress of delamination.

7. Conclusions

The solution of a hemivariational inequality can be found as a substationary point of some nonconvex functional being composed of a dominating quadratic part and a nonsmooth piecewise quadratic part. We have tested the functioning of different nonconvex and nonsmooth optimization methods in the solution of a laminated composite structure under loading when the binding material between laminae obeys a nonmonotone multivalued law.

Due to the strong quadratic nature of this kind of problem, the bundle-Newton method based on the second order piecewise quadratic model proves to be superior when compared to proximal bundle method based on the first order polyhedral approximation. Bundle-Newton method was clearly faster and more reliable, and what is best, the iteration number and the function evaluations of bundle-Newton method do not depend on the dimension of the problem. The same trend can be seen also in the academic tests of [15].

Acknowledgments

This work was supported by the grants SA-34063 and SA-8583 given by the Academy of Finland and the grant No. 201/96/0918 given by the Czech Republic Grant Agency.

References

1. Auslander, A. (1987), Numerical methods for nondifferentiable convex optimization, *Mathematical Programming Study* 30: 102–126.
2. Bihain, A. (1984), Optimization of upper semidifferentiable functions, *Journal of Optimization Theory and Applications* 4: 545–568.
3. Bonnans, J.F., Gilbert, J.C., Lemaréchal, C. and Sagastizábal, C. (1995), A family of variable metric proximal methods, *Mathematical Programming* 68: 15–47.
4. Cheney, E.W. and Golstein, A.A. (1959), Newton’s method for convex programming and Tchebycheff approximation, *Numerische Mathematik* 1: 253–268.

5. Clarke, F.H. (1983), *Optimization and Nonsmooth Analysis*. New York: J. Wiley & Sons.
6. Gaudioso, M. and Monaco, M.F. (1992), Variants to the cutting plane approach for convex nondifferentiable optimization, *Optimization* 25: 65–75.
7. Kelley, J.E. (1960), The cutting plane method for solving convex programs, *SIAM J.* 8: 703–712.
8. Kiwiel, K.C. (1985), *Methods of Descent for Nondifferentiable Optimization* (Lecture Notes in Mathematics 1133). Berlin/New York: Springer Verlag.
9. Kiwiel, K.C. (1986), A method for solving certain quadratic programming problems arising in nonsmooth optimization, *IMA Journal of Numerical Analysis* 6: 137–152.
10. Kiwiel, K.C. (1990), Proximity control in bundle methods for convex nondifferentiable optimization, *Mathematical Programming* 46: 105–122.
11. Lemaréchal, C. (1978), *Nonsmooth Optimization and Descent Methods*. IIASA-report, Laxemburg, Austria.
12. Lemaréchal, C. (1982), Numerical experiments in nonsmooth optimization, in E.A. Nurminski, (ed.), *Progress in Nondifferentiable Optimization*, (pp. 61–84). IIASA-report, Laxemburg, Austria.
13. Lemaréchal, C. and Sagastizábal, C. (1994), An approach to variable metric bundle methods, in J. Henry and J. P. Yvor (eds.), *Lecture Notes in Control and Information Sciences*, vol. 197, pp. 144–162. Berlin/New York: Springer-Verlag.
14. Lukšan, L. (1984), Dual method for solving a special problem of quadratic programming as a subproblem at linearly constrained nonlinear minmax approximation, *Kybernetika* 20: 445–457.
15. Lukšan, L. and Vlček, J. (1995), *A Bundle-Newton Method for Nonsmooth Unconstrained Minimization*. Technical report, No. 654, Institute of Computer Science, Academy of Sciences of the Czech Republic.
16. Mäkelä, M.M. and Neittanmäki, P. (1992), *Nonsmooth Optimization. Analysis and Algorithms with Applications to Optimal Control*. World Scientific, Singapore.
17. Miettinen, M. and Haslinger, J. (1995), Approximation of non-monotone multivalued differential inclusions, *IMA Journal of Numerical Analysis* 15: 475–503.
18. Miettinen, M. and Haslinger, J. (1997), Finite element approximation of vector–valued hemivariational inequalities, *Journal of Global Optimization* 10: 17–35.
19. Miettinen, M., Mäkelä, M.M. and Haslinger, J. (1995), On numerical solution of hemivariational inequalities by nonsmooth optimization methods, *Journal of Global Optimization* 6: 401–425.
20. Mifflin, R. (1996), A Quasi-second-order proximal bundle algorithm, *Mathematical Programming* 73: 51–72.
21. Mistakidis, E.S., Baniotopoulos, C.C. and Panagiotopoulos, P.D. (1995), On the numerical treatment of the delamination problem in laminated composites under cleavage loading, *Composite Structures* 30: 453–466.
22. Naniewicz, Z. and Panagiotopoulos, P.D. (1995), *Mathematical Theory of Hemivariational Inequalities and Applications*. Basel/New York: Marcel Dekker.
23. Panagiotopoulos, P.D. (1993), *Hemivariational Inequalities*. Berlin/New York: Springer Verlag.
24. Rockafellar, R.T. (1976), Monotone operators and the proximal point algorithm, *SIAM Journal on Optimal Control and Optimization* 14: 877–898.
25. Schramm, H. and Zowe, J. (1992), A version of the bundle idea for minimizing a nonsmooth functions: Conceptual idea, convergence analysis, numerical results, *SIAM Journal on Optimization* 2: 121–152.
26. Tzaferopoulos, M. Ap., Mistakidis, E.S., Bisbos, C.D. and Panagiotopoulos, P.D. (1995), Comparison of two methods for the solution of a class of nonconvex energy problems using convex minimization algorithms, *Computers and Structures* 57(6): 959–971.