Journal of Civil Engineering and Architecture 20 (2026) 322-339
doi: 10.17265/1934-7359/2026.08.002

Direct and First Inverse Problems of Genetic Identification of Materials at a Fixed Smoothness Class

P. A. Belov[1] and Elliot Weiss[2]

1. Institute of Applied Mechanics, Russian Academy of Sciences, Moscow 125040, Russia

2. Independent Researcher, USA

Abstract: Within the framework of the proposed genetics of elastoplastic materials, a mathematical formulation of the genetic identification of materials is developed on the basis of the correspondence between the experimental deformation curve, the governing differential operator, and the space of its fundamental solutions. A fixed smoothness class C^k uniquely determines the order of the governing operator, the number of second-order factors, and the dimension of the space of fundamental solutions. A canonical interpolation curve is introduced as the unique geometric reference associated with the experimental sample and the selected smoothness class; it is independent of the mechanical parameters of the material. Two complementary bounds are established. First, every proper mechanical model curve that does not coincide with the canonical interpolation curve has a strictly nonzero approximation error relative to it. Second, the generalized Reinsch construction implies that a one-segment canonical interpolant of class C^k is a polynomial of degree at most 2k+1; therefore, any finite sample of N points admits one-segment canonical interpolation for a sufficiently high class satisfying $\boldsymbol{k} \geq \left\lceil \frac{\boldsymbol{N-2}}{\boldsymbol{2}} \right\rceil$. A third result establishes a class-dependent upper bound on the number of identifiable segments: for a fixed sample size, the continuous segmentation capacity decreases strictly with smoothness class, whereas the maximum integer number of segments decreases stepwise. These results separate the class-dependent geometric limits of representation from measurement accuracy, which enters only at the stage of inverse identification. The direct and first inverse problems are then formulated for a fixed smoothness class and segment architecture.

Key words: Genetics of materials, genetic identification, direct problem, first inverse problem, canonical interpolation curve, smoothness class, factorization of a differential operator, fundamental solutions.

1. Introduction

Inverse problems occupy a central place in continuum mechanics, mathematical physics, and material-identification theory. Their general objective is to recover unknown characteristics of an object from observations. The classical concept of a well-posed mathematical problem was formulated by J. Hadamard: a solution must exist, be unique, and depend continuously on the input data [1]. Stability is especially important for inverse problems because small measurement errors may cause substantial changes in the recovered characteristics. The mathematical foundations for treating ill-posed problems were established by A. N. Tikhonov, V. K. Ivanov, and M. M. Lavrentiev and were subsequently developed within the modern theory of regularization [2-5]. The contemporary understanding of an inverse problem includes not only the recovery of unknown functions or coefficients, but also the determination of internal inhomogeneities, sources, boundaries, and operator structures from indirect data [6].

The formulation proposed in the present paper continues the authors' research program. The principal concepts of the genetics of materials were introduced in Belov [7]. A structural theory of deformation models was developed in Belov and Golovina [8], where different models were shown to admit a natural

Corresponding author: Petr A. Belov, D.Sc., Chief Research Scientist, Institute of Applied Mechanics, Russian Academy of Sciences, research fields: mechanics of materials, constitutive modeling, inverse problems, materials genetics.

interpretation as different architectures of governing operators. The present paper advances this line of research by formulating the first inverse problem of genetic identification of materials.

In material mechanics, the most common formulation is parametric: the constitutive relations and the structure of the differential operator are prescribed in advance, while the experiment is used to determine a finite set of material parameters. This approach is effective as long as the selected model family provides an adequate description of the material under investigation. If the operator structure is inconsistent with the internal architecture of the material, minimization of the residual can identify only the best representative of an incorrect family. Even accurately determined parameters cannot eliminate a structural error. Consequently, the minimum attainable residual reflects simultaneously the quality of optimization, the accuracy of the experiment, and the compatibility of the selected model architecture with the observed response.

The genetic approach changes the primary object of identification. Instead of a prescribed set of material coefficients, one considers the governing operator and the space of fundamental solutions generated by it. Factorization of the operator into primary second-order operators defines the allelic architecture of the model, whereas further factorization into linear operators defines its genes. The characteristic numbers of these operators determine the form of the fundamental solutions, while the coefficients of the linear expansion in those solutions describe the particular phenotypic manifestation of the material under prescribed loading conditions. This distinction separates stable elements of the genetic structure from coefficients associated with a particular experiment.

A fundamental parameter of the model is the smoothness class C^k. It determines the number of derivatives entering the variational formulation, the order $2(k+1)$ of the governing differential operator, the number of primary second-order operators, and the dimension of the space of fundamental solutions. Once C^k is fixed, the structural dimension of the local model is fixed, whereas the characteristic numbers, expansion coefficients, and factorization parameters remain unknown. The number of segments specifies the global construction of the curve: each segment is a new linear combination of the same space of fundamental solutions. Thus, the smoothness class and the number of segments describe two different kinds of model complexity and must not be merged into a single optimization variable.

Before formulating the inverse problem, a continuous object must be defined against which the mechanical model is evaluated. An experimental diagram is given by a finite sample of points and does not by itself determine a unique function between the measurement nodes. In the present work, this object is the canonical interpolation curve for a fixed smoothness class and prescribed segmentation. It is constructed from the geometry of the experimental sample by using the complete spectrum of boundary, contact, and continuity conditions inherited from the variational formulation. The canonical curve is not a mechanical model of the material; it provides a common geometric reference against which all mechanical models of the selected class are compared.

According to this logic, the direct problem is formulated for a known governing-operator structure, known characteristic numbers, and prescribed boundary and contact conditions. Its result is a model deformation curve represented as a piecewise linear combination of fundamental solutions. The first inverse problem follows the same trace in the opposite direction. From the experimental sample and the canonical interpolation curve, with the smoothness class and segment architecture fixed, one must recover the characteristic numbers, fundamental solutions, phenotypic coefficients, factorization, and complete governing operator. Therefore, reconstruction of the genetic structure and recovery of the operator are not two independent inverse problems, but two consecutive levels of one first inverse problem.

Fixing the smoothness class is methodologically important because it prevents implicit mixing of parameter identification within a selected architecture with selection of the architecture itself. The first inverse problem asks which operator and which genetic structure correspond to the experiment within a prescribed class C^k and for a prescribed number of segments. The same experiment may, however, admit solutions of the first inverse problem for several pairs “smoothness class-number of segments”. Comparing these solutions and selecting an optimal architecture require construction of a separate model space. This task is the subject of the next paper in the series and is indicated here only as a natural continuation of the present formulation.

The purpose of this paper is to establish a mathematical foundation for the first inverse problem of genetic identification of materials at a fixed smoothness class. To achieve this purpose, it is necessary to formulate the direct problem and establish the relation between the variational form, the governing operator, and the complete spectrum of conditions; define the canonical interpolation curve and examine the conditions of its existence and uniqueness; describe factorization of the governing operator and the corresponding genetic structure; formulate the first inverse problem as recovery of the characteristic numbers, fundamental solutions, factorization, and operator from experimental data; and determine the boundary of applicability of this problem and the transition to selection of an optimal model architecture.

The remainder of the paper is organized as follows. Section 2 formulates the direct problem: a non-integrable variational form is introduced, the governing operator and natural conditions are derived, the canonical interpolation curve is constructed, and operator factorization is considered. The existence and uniqueness of the canonical interpolation curve are then postulated, after which a rigorous formulation of the first inverse problem is given for a fixed smoothness class. Particular attention is paid to the fundamental distinction between reconstruction within a fixed model architecture and selection among alternative architectures. The concluding part identifies the conditions of existence, uniqueness, and stability of genetic identification and defines the boundary beyond which the problem of selecting an optimal model architecture arises.

2. Formulation of the Direct Problem

Formulation of the first inverse problem requires prior definition of the direct mapping that relates the structure of the mathematical model to the observed deformation curve. In the present theory, this mapping is constructed not from a preselected analytical approximation but from a non-integrable variational form. Its integral part generates the governing differential equation, the boundary terms determine the complete spectrum of natural boundary and contact conditions, and factorization of the governing operator defines the space of fundamental solutions. For known characteristic numbers and expansion coefficients, these data uniquely determine the model curve within a fixed smoothness class and segment architecture. The present section establishes a closed direct problem whose inverse is subsequently used for genetic identification of the material.

2.1 Non-integrable Variational Form and Governing Operator

Consider a normalized deformation curve $\sigma = \sigma(\varepsilon)$ defined on the interval $0 \le \varepsilon \le 1$, where ε and σ denote normalized strain and stress, respectively. For a model of smoothness class C^k, introduce the set of kinematic variables comprising the function σ and its derivatives up to order $m = k + 1$:

$$q_i = \frac{d^i\sigma}{d\varepsilon^i}, i = 1, \dots, m \qquad (1)$$

The number of derivatives employed and the maximum order m are uniquely related to the smoothness class:

$$m = k + 1 \qquad (2)$$

The reversible part of the model is defined by the quadratic form

$$J_I = \int_0^1 A_{ij} q_i q_j d\varepsilon \quad (3)$$

The quadratic form (3) defines the reversible part of the variational model, where summation over repeated indices is implied and the coefficient matrix A_{ij} of the reversible part is symmetric:

$$A_{ij} = A_{ji} \quad (4)$$

The symmetry condition (4) ensures the potential character of the reversible part.

The irreversible channels of the material are described by the antisymmetric variational form

$$\delta D = \int_0^1 B_{ij}\left(q_i \delta q_j - q_j \delta q_i\right) d\varepsilon \quad (5)$$

where the coefficients of the antisymmetric part satisfy the condition

$$B_{ij} = -B_{ji} \quad (6)$$

Relation (6) expresses the antisymmetry of the coefficients of the dissipative part.

In general, the antisymmetric part (5) is not the variation of a scalar functional. Therefore, a complete description of an elastoplastic material cannot be reduced to stationarity of a classical potential functional and requires stationarity of a non-integrable variational form. In general form, it is written as

$$\delta W = \delta J_D + \delta J_I + \delta D \quad (7)$$

where the first term denotes the variation of the discrete part of the objective functional associated with the experimental sample; the second term is the variation of the reversible integral part; and the third term is the non-integrable dissipative part describing the channels of irreversible deformation. The stationarity condition is

$$\delta W = 0 \quad (8)$$

The stationarity condition (8) is the starting point for the subsequent derivation of the governing operator.

Successive integration by parts of the variational form yields two interrelated results. The integral term generates the generalized Euler-Lagrange equation, whereas the boundary terms determine the complete spectrum of natural boundary and contact conditions. The governing operator and the conditions closing the boundary-value problem therefore have a common variational origin.

For a model of class C^k containing derivatives up to order $m = k + 1$, the highest-order quadratic term of the variational form contains the derivative $\sigma^{(m)}$. After m integrations by parts, a derivative of σ of order 2m appears. Hence, the governing equation has the form

$$L_{2m[\sigma]} = 0 \quad (9)$$

Eq. (9) gives the general form of the governing differential equation of the model, where the linear differential operator of order 2m is written as

$$L_{2m} = a_{2m} q_m^{(m)} + \cdots + a_2\, q_1^{(1)} \quad (10)$$

Operator representation (10) makes the structure of the governing operator explicit.

Thus, the smoothness class uniquely determines the order of the governing operator:

$$ord\, L_{2m} = 2m = 2(k + 1) \quad (11)$$

Relation (2) directly establishes the connection between the smoothness class and the maximum derivative order.

Relation (11) establishes a one-to-one correspondence between the smoothness class and the order of the governing operator.

Increasing the smoothness class by one increases both the order of the governing equation and the dimension of its space of fundamental solutions by two. At the same time, the number of independent conditions required for a unique construction of the model curve also increases. Thus, fixing the class C^k is not an external choice of a smoothing procedure, but a specification of the structural dimension of the local model.

To determine the complete spectrum of natural conditions, consider the boundary terms arising from successive integration by parts of the variational form. For an arbitrary model of order 2m, represent the integral part of the first variation as

$$\delta J_I + \delta D = \int_0^1 \sum_{j=0}^{m} F_j(\varepsilon)\, \delta\sigma^{(j)}\, d\varepsilon \quad (12)$$

Representation (12) is the starting form for the subsequent integration by parts, where the displayed quantity is the coefficient multiplying the variation of the j-th derivative. For the quadratic non-integrable variational form considered here, these coefficients are expressed through the matrices A_{ij} and B_{ij} and the derivatives of $\sigma(\varepsilon)$. Under the adopted summation convention, they are written as

$$F_j = 2 \sum_{j=0}^{m} (A_{ij} + B_{ij})\, \sigma^{(i)}$$

After integrating each term by parts j times, the first variation takes the form

$$\delta J_I + \delta D = \int_0^1 \sum_{j=0}^{m} (-1)^j F_j^j(\varepsilon)\delta\sigma\, d\varepsilon + \sum_{j=0}^{m-1} P_r \delta\sigma^r \tag{13}$$

The integral term in Eq. (13) generates the generalized Euler-Lagrange equation. The quantities P_r are static multipliers conjugate to the variations $\delta\sigma^{(r)}$. For an arbitrary model, they have the form

$$P_r = \sum_{j=r+1}^{m} (-1)^{j-r-1} F_j^{(j-r-1)}, r = 0, \dots, m-1 \tag{13.1}$$

The sequence of static multipliers begins with

$$P_{(m-1)} = F_m$$

$$P_{(m-2)} = F_{(m-1)} - F_m^{(1)}$$

$$P_{(m-3)} = F_{(m-2)} - F_{(m-1)}^{(1)} + F_m^{(2)}$$

$$P_0 = \sum_{j=1}^{m} (-1)^{j-1} F_j^{(j-1)}$$

Thus, each kinematic parameter $\sigma, \sigma', \dots, \sigma^{(m-1)}$ has a conjugate static multiplier $P_0, P_1, \dots, P_{m-1}$. If the value of $\sigma^{(r)}$ is prescribed at an endpoint, then $\delta\sigma^{(r)} = 0$. If it is not prescribed, the corresponding variation remains arbitrary, and stationarity requires the natural condition $P_r = 0$.

At the left endpoint of the domain, the kinematic conditions are prescribed as

$$\sigma(0) = 0, \sigma''(0) = 0$$

Therefore, $\delta\sigma(0) = 0$ and $\delta\sigma''(0) = 0$. The variation $\delta\sigma'(0)$ remains free, and the conjugate static multiplier must consequently vanish:

$$P_1(0) = 0$$

Variations of derivatives of order higher than two also remain free. Therefore, the following natural conditions hold at the left endpoint:

$$P_r(0) = 0, r = 3, \dots, m-1$$

At the right endpoint of the domain, the kinematic conditions are prescribed as

$$\sigma(1) = 1, \sigma'(1) = 0$$

Therefore, $\delta\sigma(1) = 0$ and $\delta\sigma'(1) = 0$. All variations of derivatives beginning with the second order remain free, and the following natural conditions hold at the right endpoint:

$$P_r(1) = 0, r = 2, \dots, m-1$$

If the indicated index range is empty for a low-order model, the corresponding conditions are absent. Thus, the direct boundary-value problem for an arbitrary model of order 2m comprises the generalized Euler-Lagrange equation, four prescribed kinematic conditions, and a set of homogeneous natural conditions expressed through the static multipliers P_r.

The canonical interpolation curve arises as a special geometric representative of the general variational model. In its construction, the contributions of all lower-order derivatives are eliminated:

$$F_0 = F_1 = \cdots = F_{(m-1)} = 0$$

The only nonzero coefficient is the highest-order coefficient

$$F_m = a_{2m}\, \sigma^{(m)}, a_{2m} \neq 0$$

Therefore, the generalized Euler-Lagrange equation becomes

$$(-1)^m a_{2m} \sigma^{(2m)} = 0$$

Since the coefficient is nonzero, this equation is equivalent to

$$\sigma^{(2m)} = 0$$

The static multipliers for the canonical interpolation curve reduce to

$P_r = (-1)^{(m-r-1)} a_{2m} \sigma^{(2m-r-1)}, r = 0, \dots, m-1$

In particular,

$$P_{(m-1)} = a_{2m}\,\sigma^{(m)}$$
$$P_{(m-2)} = -a_{2m}\,\sigma^{(m+1)}$$
$$P_{(m-3)} = a_{2m}\sigma^{(m+2)}$$
$$P_0 = (-1)^{(m-1)}\,a_{2m}\,\sigma^{(2m-1)}$$

Substitution of these expressions into the general boundary conditions yields the boundary-value problem for the canonical interpolation curve:

$$\sigma(0) = 0, \sigma^{(2)}(0) = 0, \sigma^{(2m-2)}(0) = 0, \sigma^{(2m-r-1)}(0) = 0 \; (r = 3, \dots, m-1) \tag{14}$$
$$\sigma(1) = 1,\; \sigma^{(1)}(1) = 0, \sigma^{(2m-r-1)}(1) = 0 \; (r = 2, \dots, m-1)$$

Eq. (14) defines the complete boundary-value problem for the canonical interpolation curve.

The common multiplier does not affect either the solution of the differential equation or the formulation of the boundary conditions for the interpolation curve. Consequently, it does not affect the shape of the canonical interpolation curve and determines only the relative weight of the integral part of the objective functional with respect to the discrete residual. This separates geometric reconstruction of the experimental sample from the mechanical parameters of the material.

For an r-segment canonical interpolation curve (spline), the external boundary conditions are supplemented by continuity conditions between adjacent segments:

$$\sigma_r^{(j)}(\varepsilon_r) = \sigma_{(r+1)}^{(j)}, j = 0, \dots, k$$

These conditions ensure global smoothness of class C^k and unique assembly of the segments, but they do not alter the structure of the external boundary-value problem and do not introduce the coefficient into the geometry of the interpolation curve. Thus, the choice of smoothness class determines the operator order 2m, the dimension of the space of fundamental solutions, the spectrum of static multipliers, and the admissible manner of joining the segments. The resulting boundary-value formulation provides the immediate basis for defining the canonical interpolation curve in the next subsection.

2.2 Generalized Reinsch Principle and the Canonical Interpolation Curve

An experiment directly provides not a continuous diagram but a finite sample of normalized points

$$E = \{(\varepsilon_n, \sigma_n)\}_{n=1}^{n=N} \tag{15}$$

Eq. (15) defines the experimental sample used to construct the canonical interpolation curve.

To pass from a discrete sample to a continuous object, an interpolation curve of a prescribed smoothness class must be constructed. The classical Reinsch principle relates the closeness of the sought function to the experimental data to a constraint on its curvature. In the present theory, this principle is generalized to an arbitrary variational model of class C^k and the corresponding operator of order 2m = 2(k + 1). Instead of a single curvature measure, the full variational structure is used, including derivatives up to order m and the associated spectrum of boundary, contact, and continuity conditions.

A rigorous proof of the existence and uniqueness of such a curve is a separate problem in the general theory of variational interpolation. For the formulation of the first inverse problem, however, it is sufficient to postulate the existence of a common continuous geometric reference that is the same for all mechanical models of the selected smoothness class. We therefore adopt the following statement.

Postulate of existence and uniqueness of the canonical interpolation curve: "For a prescribed sample of experimental points, a fixed smoothness class C^k, a selected segmentation, and the corresponding complete spectrum of boundary, contact, and continuity conditions, there exists a unique interpolation curve of class k:

$$I(\varepsilon, k, s) \tag{16}$$

where s is the number of segments". In the present work, the existence and uniqueness of this curve are accepted as a consequence of the generalized Reinsch principle.

Postulate (16) establishes the existence and uniqueness of the canonical interpolation curve.

The canonical interpolation curve is determined solely by the geometry of the experimental sample, the selected smoothness class, the segmentation, and the continuity conditions. Since the coefficient cancels from both the Euler-Lagrange equation and the homogeneous natural boundary conditions, construction of the curve is independent of its value. Consequently, the interpolation curve is independent of the mechanical properties of the material and is determined exclusively by the geometry of the experimental data and the selected smoothness class.

For fixed k and s, the function in (16) is the unique continuous representative of the prescribed sample within the adopted smoothness class. Therefore, different mechanical models of this class are compared neither directly with one another nor with an arbitrarily selected smoothing function, but with the same canonical geometric object—the interpolation curve.

The canonical interpolation curve is thereby introduced as an independent geometric object: it reconstructs the continuous form of the experimental diagram but does not inherit the structure or parameters of any preselected mechanical model. It is precisely this independence from the family of models being compared that allows it to serve as the common reference for the first inverse problem.

The canonical interpolation curve serves as the zero reference of the least-squares method when models of the selected smoothness class are compared. For the model curve

$$M(\varepsilon, \theta, k, s) \tag{17}$$

where θ is the parameter vector of a particular mechanical model, the residual is defined relative to the function

$$RMSE(\theta, k, s) = \left[\left(\frac{1}{N_q} \right) \Sigma_{\nu=1}^{N_q} \left(M(\hat{\varepsilon}_\nu, \theta, k, s) - I(\hat{\varepsilon}_\nu, k, s) \right)^2 \right]^{\frac{1}{2}} \tag{18}$$

Relations (17) and (18) define the model curve and the corresponding residual.

Here, the displayed set is a common control grid that may include the original experimental nodes and additional points between them. Intermediate points are necessary because, at strict interpolation nodes, the canonical function coincides with the sample, whereas differences among mechanical models appear over the entire continuous range.

Thus, the optimization problem is transferred from the discrete experimental sample to a continuous geometric object common to all mechanical models of the given smoothness class. This removes the dependence of the comparison result on an arbitrarily selected method of connecting the experimental points.

Consequently, the canonical interpolation function is not a mechanical model of a particular material. It is a canonical geometric reconstruction of the sample in the prescribed smoothness class and establishes a common zero level for evaluating all mechanical models of that class.

The result of the interpolation problem is a function determined by the sample geometry and the smoothness class. The result of the direct mechanical problem is a model deformation curve generated by the known structure of the governing operator and the complete spectrum of boundary and contact conditions. The quality of the mechanical model is characterized by the distance between these two curves.

Accordingly, comparison of the model and canonical curves requires explicit construction of the model curve. This, in turn, requires disclosure of the internal structure of the governing operator, determination of its factorization, and construction of the corresponding space of fundamental solutions. These issues are addressed in Section 2.3.

2.2.1 Greedy Segmentation as a Common Construction Principle

The greedy segmentation principle is used in the present framework for two different curve-construction problems: the canonical interpolation curve and the mechanical model curve. In both cases, the process starts from a one-segment representation and successively introduces a new segmentation node selected from the experimental sample. However, the two procedures must not be identified. For the canonical interpolation curve, the trial node is selected according to the reduction of the interpolation residual produced by the class-dependent Reinsch operator. For the mechanical model curve, the trial node is selected according to the reduction of the model residual produced by the corresponding space of fundamental solutions. Thus, the segmentation strategy is common, whereas the local curve generator and the optimized objective are different.

This distinction is essential computationally. The interpolation branch constructs the unique class-dependent geometric reference and contains no mechanical parameters. The model branch searches within the physical family generated by the fundamental solutions. A node accepted by one branch is therefore not automatically optimal for the other branch.

2.2.2 Greedy Construction of the Canonical Interpolation Curves C^0, C^1, and C^2

The following figure blocks illustrate the same local greedy construction principle for the first three smoothness classes. At every step, all admissible sample points

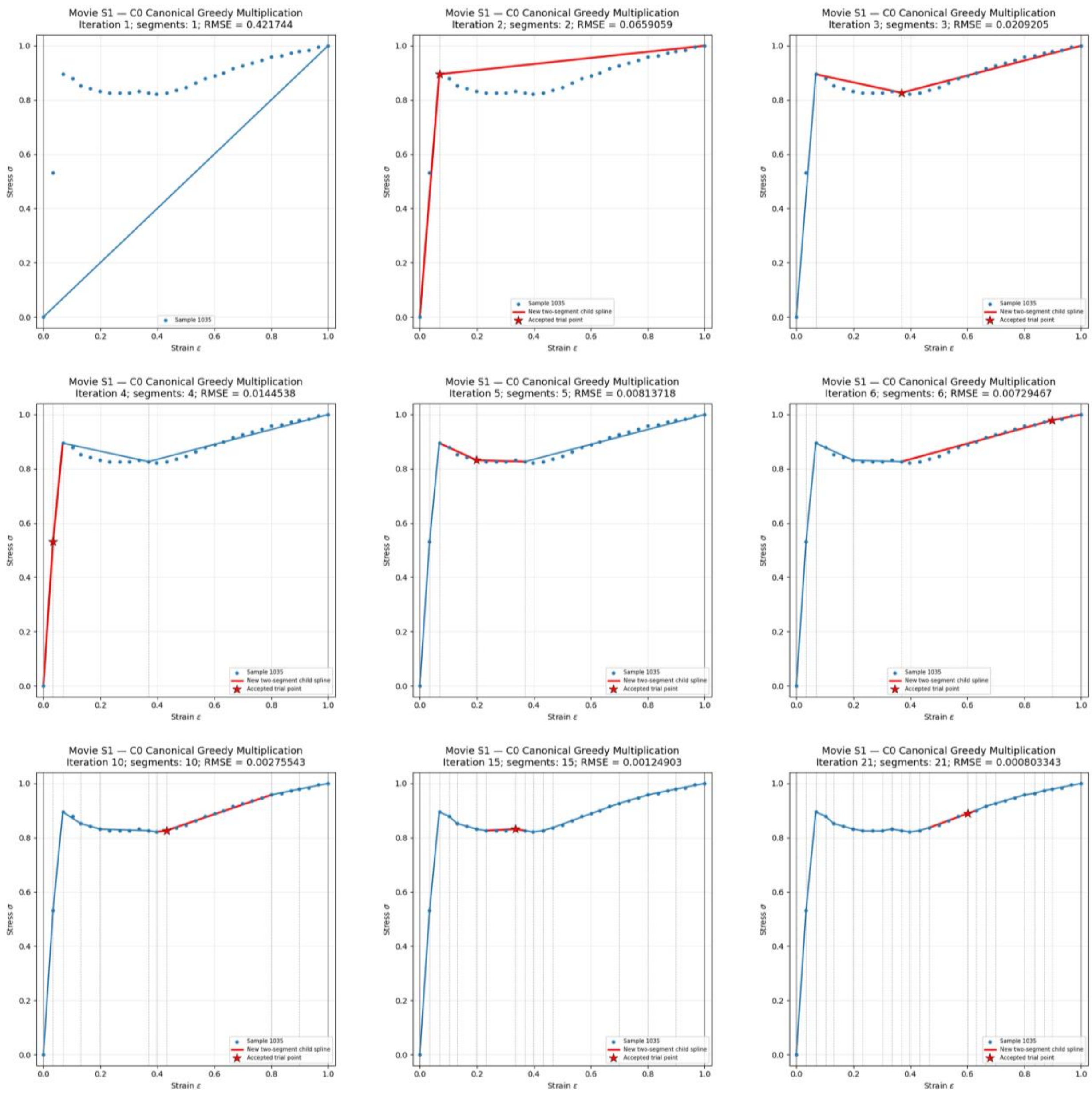

Fig. 1 Representative stages of local greedy multiplication of the canonical interpolation curve of class C^0 for material 1035.

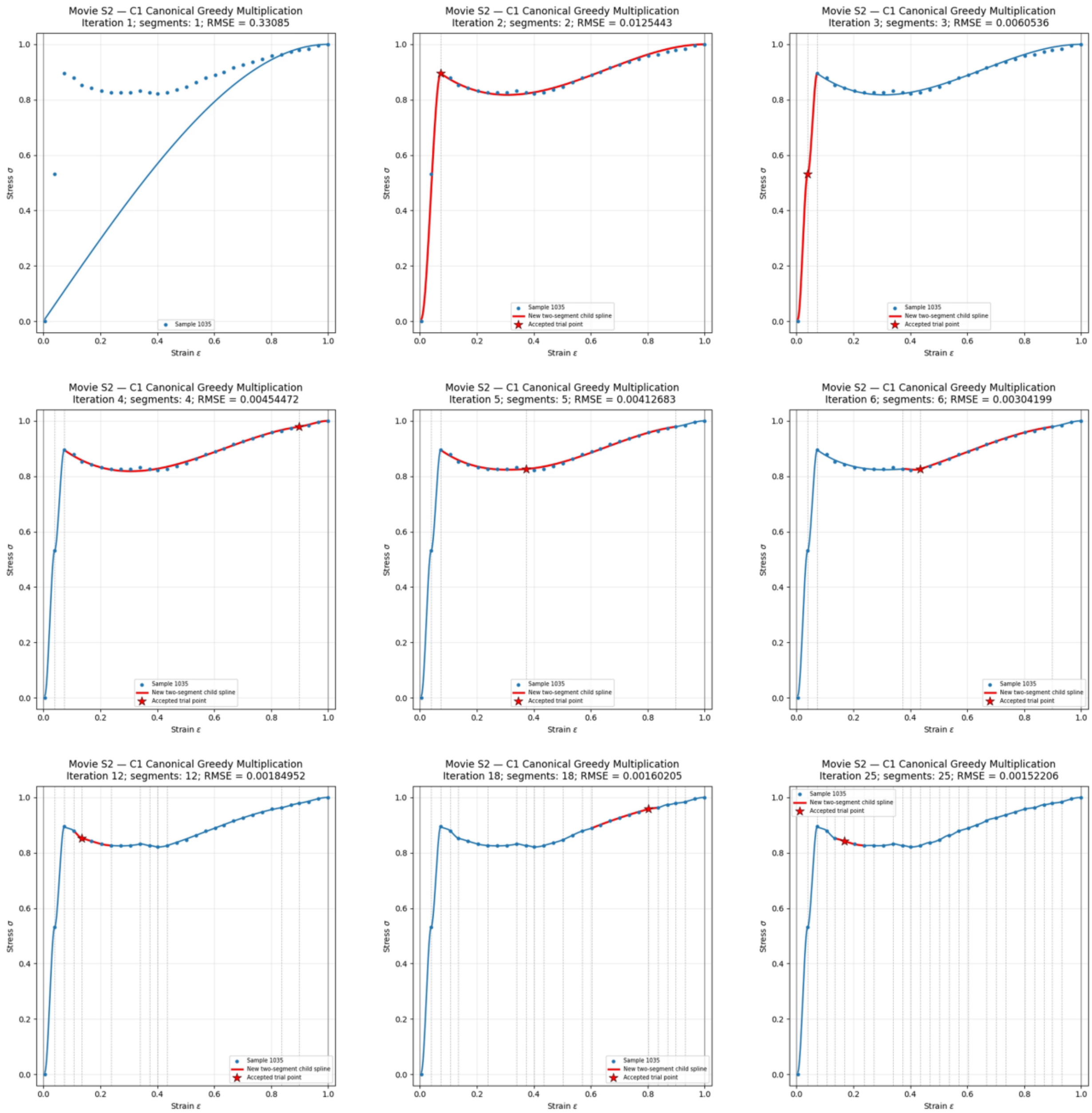

Fig. 2 Representative stages of local greedy multiplication of the canonical interpolation curve of class C¹ for material 1035.

inside each current parent segment are tested as prospective new segmentation nodes. The trial point producing the greatest decrease of the interpolation RMSE is accepted. One selected parent segment is then replaced by exactly two child segments, whereas every segment outside the selected parent interval is inherited without recalculation. Thus, one iteration increases the number of segments by exactly one and changes the curve only inside the selected parent interval. The accepted trial point is marked by a star. The newly constructed two-segment child spline is shown in red, while all inherited segments remain blue. Only the class-dependent local Reinsch interpolant changes: straight segments for C^0, cubic segments for C^1, and quintic segments for C^2.

For class C^0, the canonical local generator is linear. The interpolation curve and the parameter-free C^0 model space coincide. Each accepted point divides one selected parent line segment into two child line segments; all other segments are inherited unchanged. The process terminates at the finite-sample segmentation limit.

For class C^1, the local canonical generator is cubic. Each iteration performs a local cascade operation: one parent cubic segment is replaced by a two-segment cubic spline satisfying the inherited endpoint conditions and the C^1 contact condition at the accepted trial point. The new two-segment spline is highlighted in red, whereas the inherited segments remain blue. Unlike the mechanical C^1 model, the canonical curve contains no parameters of the two-parameter second allele.

For class C^2, the local canonical generator is quintic. One parent quintic segment is replaced by a two-segment quintic spline satisfying the inherited endpoint conditions and the C^2 contact conditions at the accepted

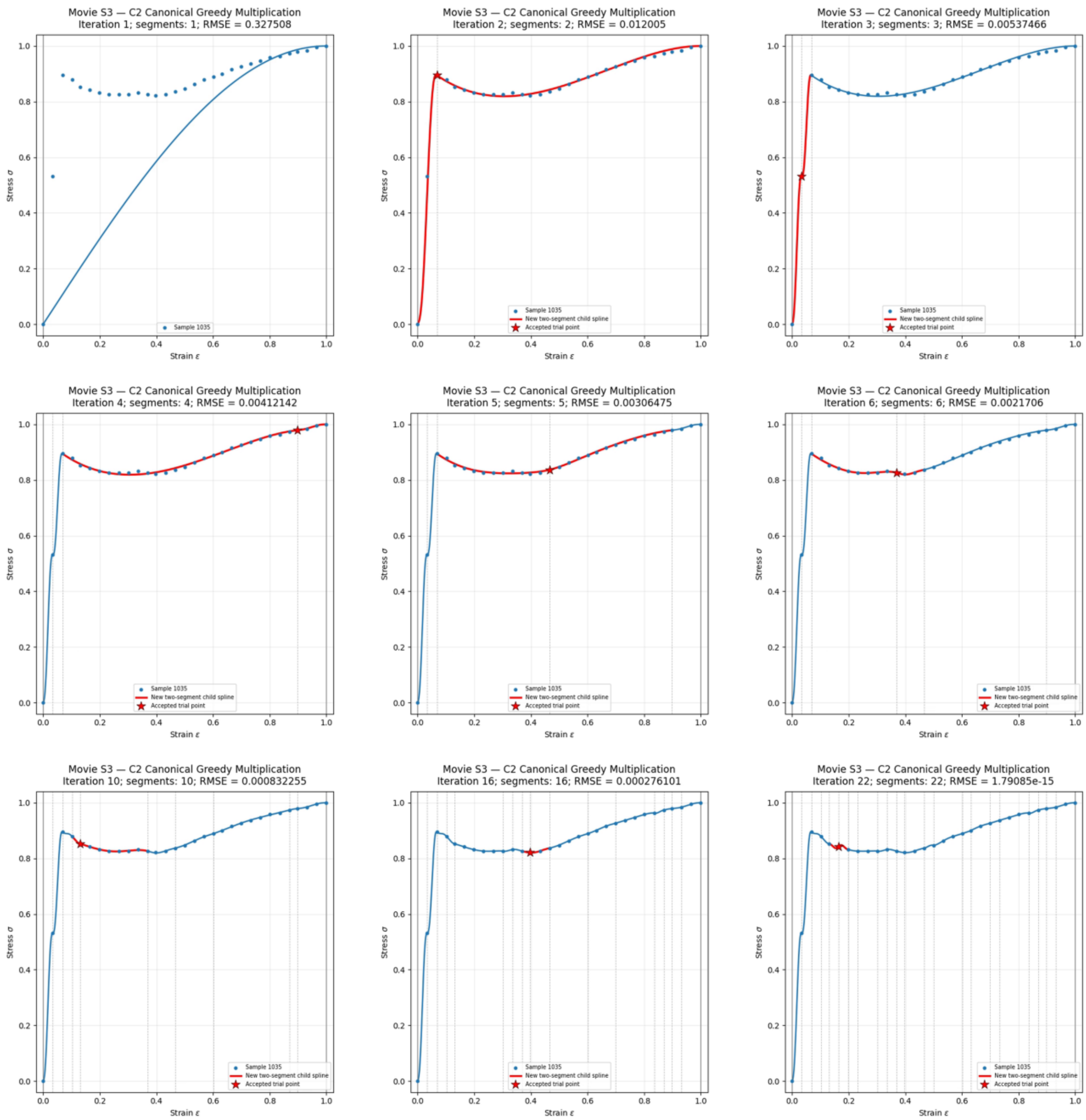

Fig. 3 Representative stages of local greedy multiplication of the canonical interpolation curve of class C^2 for material 1035.

trial point. Segments outside the parent interval are not recomputed. The accepted trial point is selected exclusively by the decrease of the C^2 interpolation residual. No inherited or newly identified characteristic exponents enter this construction; those quantities belong only to the mechanical model branch.

Figures 1-3 demonstrate that greedy multiplication is a local cascade mechanism rather than a global refitting of all previously constructed segments. Its invariant is the exact inheritance of every nonselected segment. The smoothness class determines the local interpolation operator and the corresponding contact conditions, whereas the same parent-to-two-children rule is subsequently applicable to a mechanical model curve with a different residual and a different local generator. This constructive separation prepares the lower and upper estimates established below.

2.2.3 Greedy Construction of the Mechanical Model Curves C^0, C^1, and C^2

The same local parent-to-two-children strategy is applied to mechanical model curves in Figures 4-6, but the local generator is now formed by the class-dependent

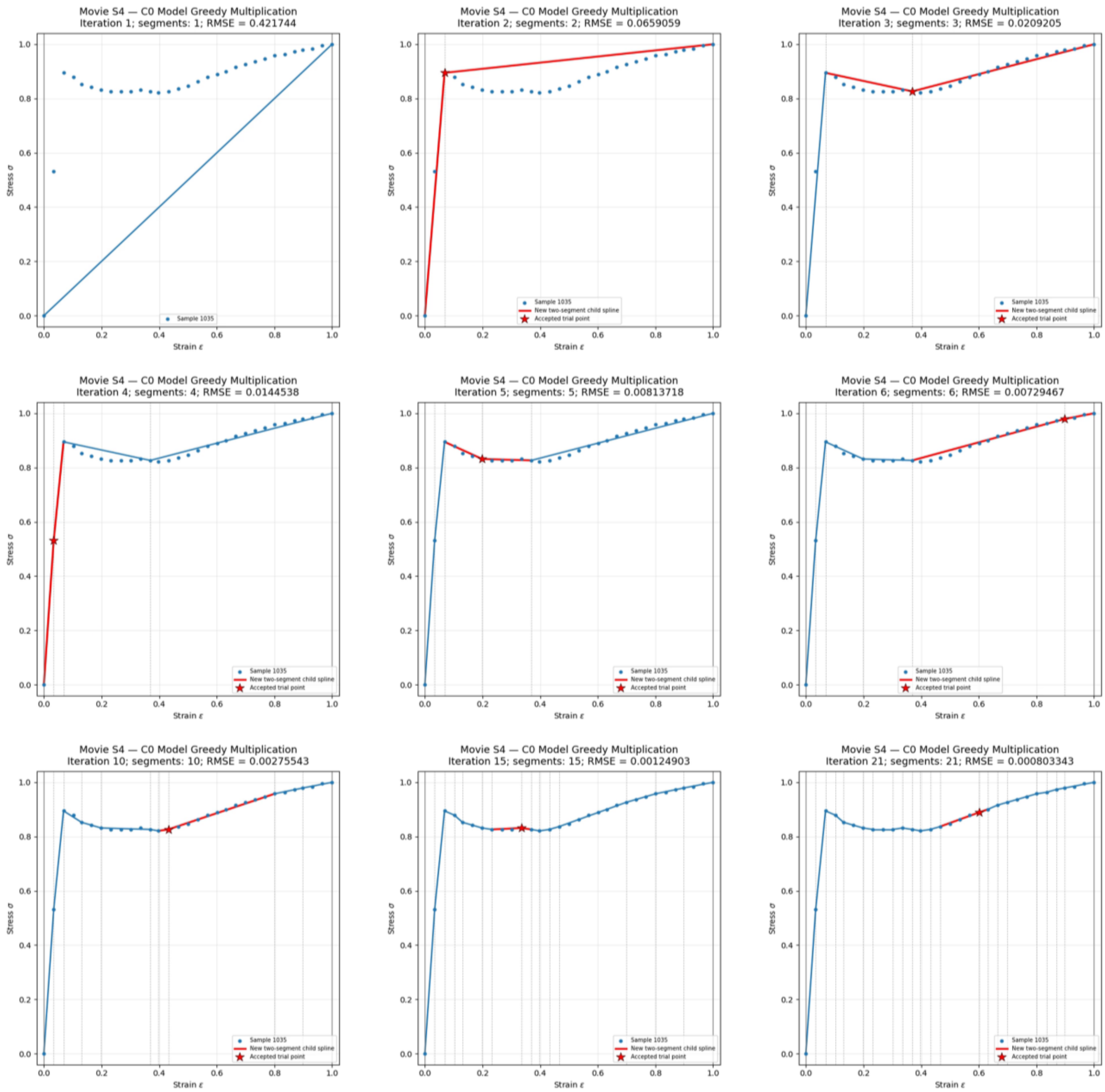

Fig. 4 Representative stages of local greedy multiplication of the mechanical model curve of class C^0 for material 1035.

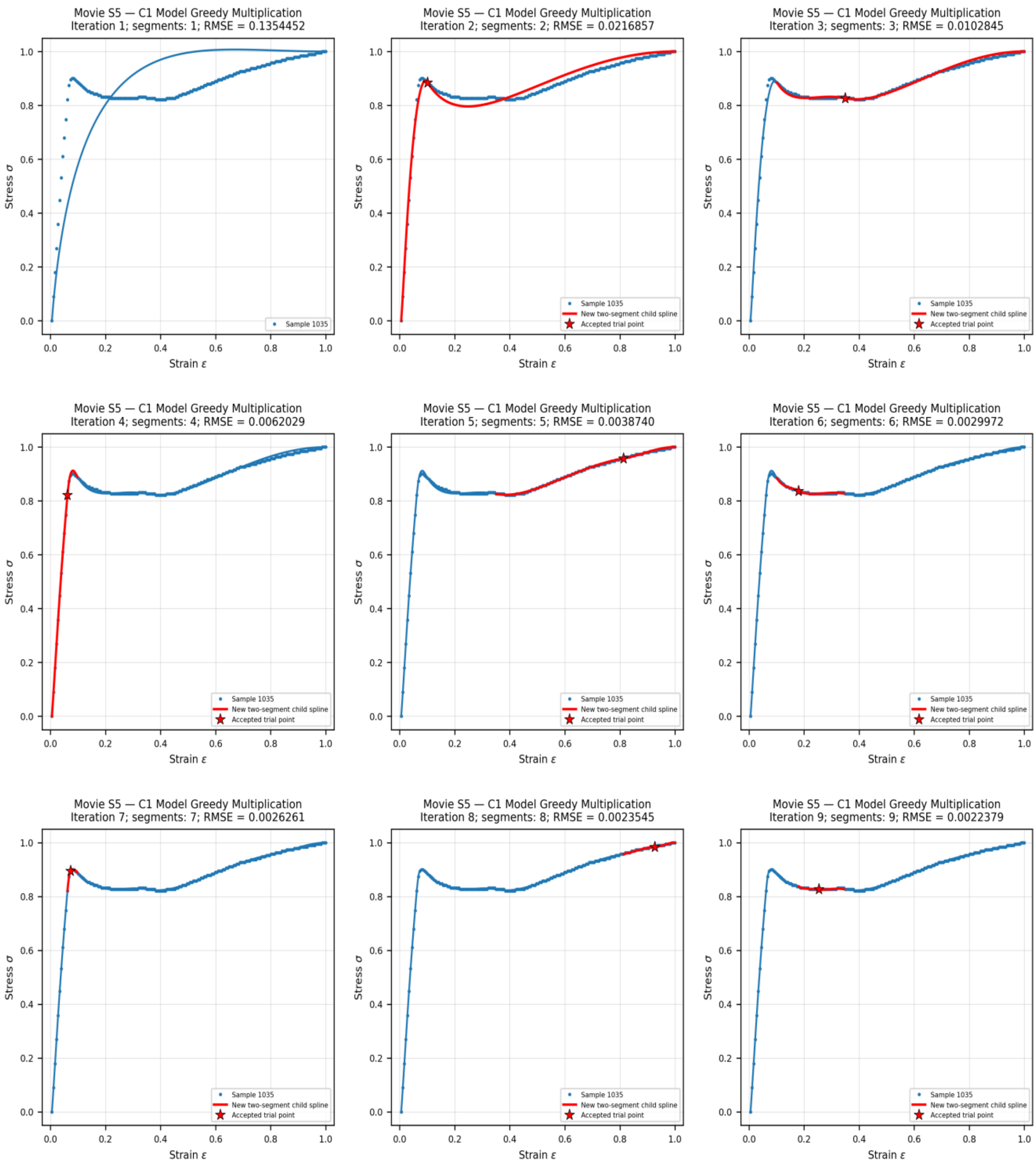

Fig. 5 **Representative stages of local greedy multiplication of the mechanical model curve of class C¹ for material 1035.**

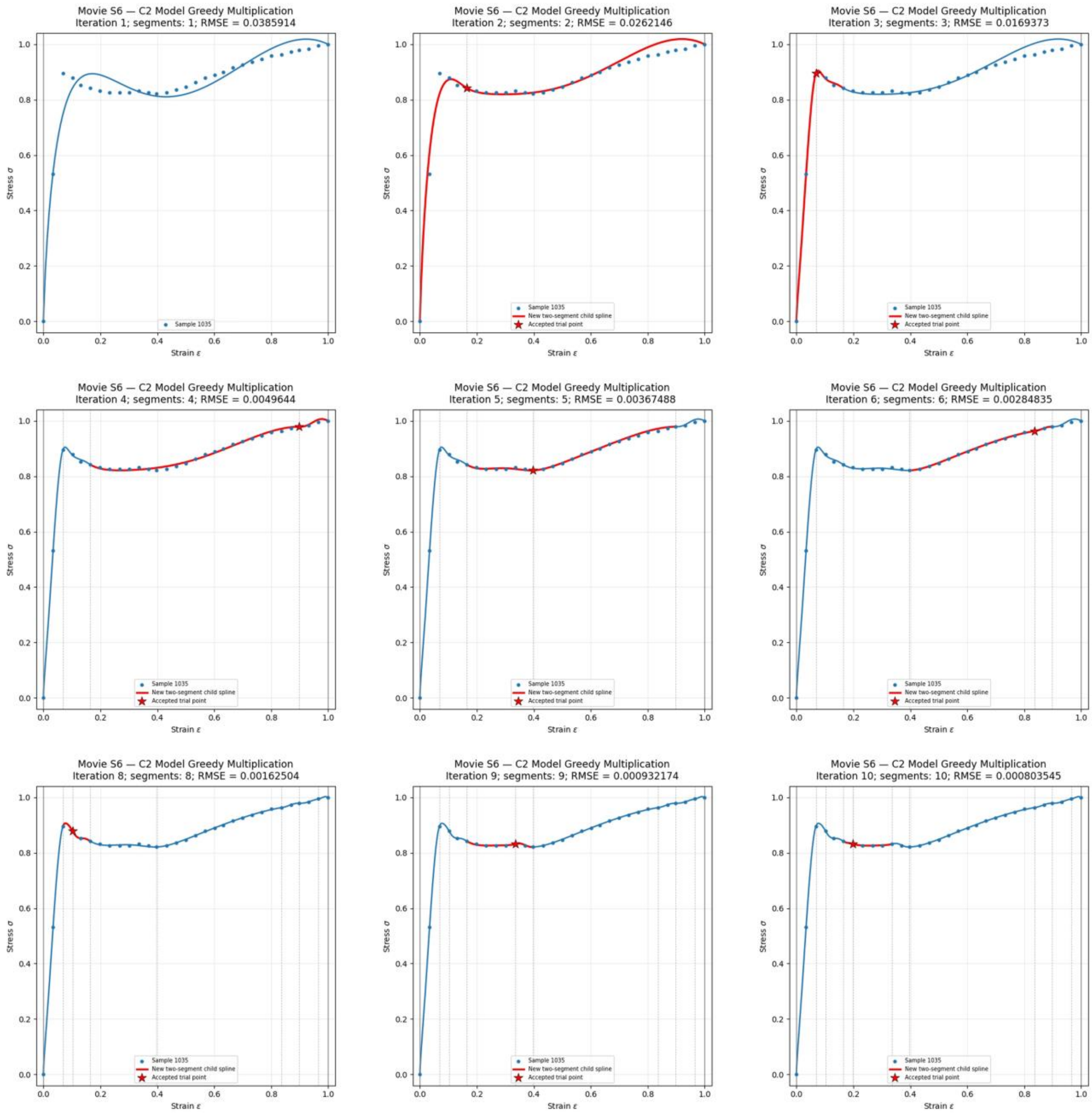

Fig. 6 Admissible stages of local greedy multiplication of the mechanical model curve of class C^2 for material 1035: iterations 1-6, 8, 9, and the maximum identifiable 10-segment state for the 31-point sample.

space of fundamental solutions, and the accepted node minimizes the model residual. The red two-segment child curve replaces only the selected parent segment; all blue segments are inherited without recalculation.

For class C^0, the model and canonical interpolation spaces coincide because the local generator contains no mechanical characteristic exponents.

For class C^1, the local model generator is defined by the two characteristic exponents of the identified second allele. Each accepted node locally replaces one parent model segment by two child segments in the same two-parameter solution space.

For class C^2, the first pair of characteristic exponents is inherited from C^1 and the additional pair is identified within C^2. For the present 31-point sample, the segmentation-capacity bound gives s_max = 10; therefore, states with 11 or more segments are excluded from the figure as mathematically underdetermined.

The four characteristic exponents determine the local model solution space, while the cascade invariant preserves every nonselected segment exactly.

2.2.4 Lower and Upper Bounds of Canonical Approximation

The canonical interpolation curve is a single class-dependent geometric object. For a prescribed experimental sample and smoothness class C^k, it is generated by the canonical Reinsch operator and does not contain mechanical parameters of the material. Measurement-system accuracy is not involved in the following estimates. At this stage, the limiting process is governed only by segmentation and by the finite number of sample points.

Let σI(ε) denote the canonical interpolation curve of class C^k and let σM(ε; θ) be a mechanical model curve of the same class, where θ is the vector of mechanical parameters. Their discrepancy is evaluated on a common control grid {εq} containing the sample nodes and, when necessary, intermediate points:

$$RMSE_{(}M/I_{)}(\theta) = [\,(1/Q)\,\Sigma(q = 1 \ldots Q)\,(\sigma M(\varepsilon q;\theta) - \sigma I(\varepsilon q))^2\,]^{1/2}.$$

Theorem 1 (nonzero error of every proper model curve).

If a mechanical model curve is, by definition, different from the canonical interpolation curve, then its error relative to the canonical curve is strictly positive.

Proof.

If $RMSE_{(}M/I_{)}$ were equal to zero, every squared difference on the control grid would be zero. Hence σM(εq;θ) = σI(εq) at every control point. For a control grid that resolves the adopted segmented representation, this equality determines the same continuous curve. Therefore σM = σI, which contradicts the definition of a proper mechanical model curve as a curve distinct from the canonical interpolation curve. Consequently:

$$\sigma M \neq \sigma I \Rightarrow RMSE_{(}M/I_{)} > 0.$$

Thus, the nonzero residual is not caused by measurement uncertainty. It follows from the geometric distinction between the model family and the unique canonical interpolation curve of the same smoothness class.

Corollary 1 (lower bound within a prescribed model family).

When the admissible parameter range is finite and the model curve depends continuously on its parameters, the least value of $RMSE_{(}M/I_{)}$ is attained. Since the canonical interpolation curve is excluded from the mechanical model family, this minimum is strictly positive:

$$RMSEmin^{(k)} = min\theta\; RMSE_{(}M/I_{)}(\theta) > 0.$$

This quantity is the theoretical lower accuracy boundary of the selected mechanical model family at the fixed class C^k and fixed segment architecture.

Theorem 2 (upper class bound for one-segment canonical interpolation).

For any finite sample of N points with distinct strain coordinates, there exists a finite smoothness class whose canonical interpolation curve can represent the entire sample by one segment.

Proof.

For class C^k, the generalized Reinsch construction uses m = k + 1 and the canonical equation contains only the highest derivative:

$$D^{2k+2}\,\sigma I = 0.$$

Therefore, on one segment the canonical fundamental solution is a polynomial of degree not exceeding 2k+1. A sample of N points is interpolated by a polynomial of degree not exceeding N-1. Hence a sufficient condition for one-segment canonical interpolation is

$$2k + 1 \geq N - 1,$$

or equivalently

$$k \geq ceil((N - 2)/2).$$

Thus, the number of sample points provides a finite upper estimate of the smoothness class required for one-segment canonical interpolation. This theorem concerns canonical interpolation curves only; it does not state that a physical model of the same class reproduces the sample exactly.

Theorem 3 (monotone decrease of segmentation capacity with smoothness class).

For a fixed sample containing $N > 2$ points, an increase in the smoothness class strictly decreases the continuous segmentation capacity of the architecture and decreases stepwise the maximum integer number of identifiable segments.

Proof.

Let the class C^k correspond to a differential operator of order 2m, where $m = k + 1$. The variational formulation supplies m independent boundary terms at each end of a segment; equivalently, a solution of the order-2m equation is uniquely specified by m independent nodal conditions at the left end and m independent nodal conditions at the right end. For a spline composed of s segments, there are s + 1 nodal points, including the two external endpoints. At each external endpoint, all m nodal conditions are fixed by the boundary-value formulation. At each of the s - 1 internal nodes, the sample fixes the value of σ, whereas the remaining m - 1 nodal quantities must be identified from the other sample points.

After the s + 1 nodal sample points have been assigned, N - (s + 1) points remain available as independent scalar conditions. Therefore, a necessary information-balance condition for identifiability is

$$N-(s+1) \geq (m-1)(s-1) \tag{18a}$$

Solving inequality (18a) for s and using $m = k + 1$ gives the continuous upper estimate

$$s \leq \bar{s}(N,k) = 1 + \frac{N-2}{k+1} \tag{18b}$$

Since the number of segments is integer, the maximum identifiable number of segments is

$$s_{\max}(N,k) = 1 + \mathrm{floor}\left(\frac{N-2}{k+1}\right) \tag{18c}$$

For fixed $N > 2$, differentiation of the continuous capacity with respect to k yields

$$\frac{\partial \bar{s}}{\partial k} = -\frac{N-2}{(k+1)^2} < 0 \tag{18d}$$

Hence, the continuous segmentation capacity strictly decreases with increasing smoothness class. Applying the floor operation to a strictly decreasing continuous bound produces a nonincreasing integer-valued function. Therefore, the maximum integer number of identifiable segments decreases stepwise as the smoothness class increases. For any $k_2 > k_1$, one has $\bar{s}(N,k_2) < \bar{s}(N,k_1)$ and $s_{\max}(N,k_2) \leq s_{\max}(N,k_1)$. The theorem is proved.

For the first three classes, relation (18c) gives $s_{\max}(N,0) = N - 1$, $s_{\max}(N,1) = \mathrm{floor}(N/2)$, and $s_{\max}(N,2) = \mathrm{floor}((N + 1)/3)$. Thus, for a fixed sample, admissible architectures occupy a class-dependent bounded region rather than a rectangular product of arbitrary smoothness classes and arbitrary segment numbers.

The established bounds describe complementary limits of the architecture space. At a fixed class, every proper mechanical model remains at a strictly positive distance from its unique canonical interpolation curve. When the class is increased, a finite upper class exists at which the complete finite sample can be represented by a single canonical interpolation segment. At the same time, for a fixed sample size, increasing the smoothness class strictly reduces the continuous segmentation capacity and stepwise reduces the maximum integer number of identifiable segments. None of these statements uses the accuracy of the measuring system. Measurement accuracy enters only in the inverse problems, where the stability and reliability of recovered characteristic numbers and operator parameters are evaluated.

2.3 Factorization of the Governing Operator

For a fixed smoothness class C^k, the number $m = k + 1$ determines the order $2m = 2(k + 1)$ of the governing operator. Hence, the smoothness class, the number of second-order operators, and the order of the complete operator are not independent characteristics of the model.

After the governing operator has been constructed, it is factorized into a product of second-order operators, which, for brevity and by analogy with genetics, will be termed alleles:

$$L_{2m} = a_{2m} \prod_{\alpha=1}^{m} L_2^{\alpha} \tag{19}$$

Factorization (19) defines the allelic architecture of the governing operator.

Each allele has the form

$$L_2^{\alpha} = D^2 + p_{\alpha} D + q_{\alpha} \tag{20}$$

Representation (20) defines the general form of an individual allele.

Or an equivalent form after transition to the corresponding Euler operator. Factorization exists for a real linear operator of even order: its characteristic polynomial can be represented over the real field as a product of m quadratic factors.

At the next level, each allele (primary second-order operator) is factorized into a product of linear (secondary) operators, which, for brevity and by analogy with genetics, will be termed genes:

$$L_2^{\alpha} = (D - \lambda_{\alpha 1})(D - \lambda_{\alpha 2}) \tag{21}$$

The two displayed quantities are the characteristic roots of the corresponding allele. The linear operators correspond to the fundamental solutions of the governing equation. Depending on the structure of the characteristic roots, the genes may be real, complex-conjugate, or repeated.

Thus, factorization of the governing operator has a two-level structure:

$$L_{2m} \rightarrow \prod_{\alpha=1}^{m} L_2^{\alpha} \rightarrow \prod_{\alpha=1}^{m} (D - \lambda_{\alpha 1})(D - \lambda_{\alpha 2}) \tag{22}$$

The two-level scheme (22) combines the allelic and genetic structures of the operator.

The first level defines the allelic architecture of the operator. The number of alleles is completely fixed by the smoothness class:

$$m = k + 1 \tag{23}$$

Relation (23) shows that the number of alleles is completely determined by the smoothness class.

The second level defines the genetic structure of each allele through its characteristic roots. It is the values and organization of these roots that individualize a particular material within the selected smoothness class.

For a prescribed class C^k, the space of fundamental solutions has dimension

$$dim\, G_k = 2m = 2(k + 1) \tag{24}$$

Eq. (24) defines the dimension of the space of fundamental solutions.

On the r-th segment, the model curve is represented as a linear combination of the fundamental solutions:

$$M^r(\varepsilon, k, s) = \Sigma_{\beta=1}^{2m} c_{\beta}^{r} g_{\beta}(\varepsilon) \tag{25}$$

Expansion (25) represents the model curve in terms of the fundamental solutions, where the first displayed quantities are the fundamental solutions (genes), while the second are the phenotypic coefficients of the given segment. Changing the segmentation does not change the underlying solution space: each new segment is a new linear combination of the same fundamental solutions. This makes it possible to distinguish rigorously between changing the model class and changing the number of segments.

For a fixed smoothness class and a prescribed genetic structure, the direct problem consists in determining the coefficients that satisfy the external boundary conditions and the intersegment continuity conditions, and in constructing the model deformation curve.

The direct trace therefore has the form:

$$\begin{aligned} &non - integrable\ variational\ form \\ &\qquad \rightarrow\ Euler \\ &\qquad - Lagrange\ equation \\ &\qquad \rightarrow\ L2m \\ &alleles\ \rightarrow\ genes \\ &\rightarrow\ model\ deformation\ curve \end{aligned} \tag{26}$$

Trace (26) completes the formulation of the direct problem and prepares the transition to the first inverse problem.

The first inverse problem is the inverse of the direct problem in the operator sense. Whereas the direct problem maps the governing operator to the model deformation curve, the first inverse problem recovers the governing differential operator from the canonical interpolation curve constructed from the experimental sample. The immediate object of recovery is the governing differential operator. The characteristic roots, allelic structure, fundamental solutions, and expansion coefficients constitute successive levels of its reconstruction. The first inverse problem of genetic identification of materials consists in recovering the complete governing differential operator at a fixed smoothness class and segment architecture. Its input data are the experimental sample and the canonical interpolation curve constructed from it. For a fixed smoothness class and the corresponding number of alleles, one must recover the characteristic roots of the genes, the allelic structure, and the complete governing operator, including its leading coefficient. This mirror symmetry between the direct and first inverse problems ensures the unity of the entire formulation.

3. Conclusions

This paper establishes the mathematical foundation of the first inverse problem of genetic identification of materials at a fixed smoothness class and segment architecture. The central object of the study is the governing differential operator, and identification is treated as the problem of recovering this operator from the experimental deformation diagram represented by the canonical interpolation curve.

It is shown that the smoothness class uniquely determines the order of the governing operator, the number of alleles, and the dimension of the space of fundamental solutions. The canonical interpolation curve is introduced as a common geometric reference independent of the mechanical model and ensuring a consistent comparison of different model architectures.

The principal scientific result is a fundamental distinction between two different classes of inverse problems. The first inverse problem reconstructs the governing operator within a model architecture fixed in advance. Selection of the optimal pair "smoothness class-number of segments", that is, selection among alternative model architectures, does not belong to this problem and constitutes a separate second inverse problem. Thus, the proposed formulation distinguishes two fundamentally different tasks: recovery of the operator within a fixed architecture and selection of an optimal architecture in the space of admissible models.

The results provide a mathematical basis for the further development of genetic identification of materials, the construction of operator-recovery algorithms, and the subsequent investigation of the space of model architectures.

Two class-dependent bounds of approximation have also been established. The lower bound follows from the fact that every proper mechanical model curve differs from the unique canonical interpolation curve and therefore has a strictly nonzero residual relative to it. The upper class bound follows from the generalized Reinsch operator: a one-segment canonical interpolant of class C^k is a polynomial of degree at most $2k+1$, so a finite sample of N points admits one-segment canonical interpolation for $k \geq [((N-2)/2)]$. In addition, the segmentation-capacity theorem shows that, for a fixed sample size, increasing the smoothness class strictly reduces the continuous segmentation capacity and decreases the maximum integer number of identifiable segments stepwise.

These bounds concern the geometry of finite-sample representation and must not be confused with measurement accuracy. The latter becomes relevant only in the inverse identification stage, where uncertainty in the experimental data propagates into the recovered genetic and operator parameters.

References

[1] Hadamard, J. 1923. *Lectures on Cauchy's Problem in Linear Partial Differential Equations*. New Haven: Yale University Press. 316 pp.

[2] Tikhonov, A. N. 1963. "Solution of Incorrectly Formulated Problems and the Regularization Method." *Soviet Mathematics Doklady* 4: 1035-8.

[3] Ivanov, V. K. 1963. "On Ill-Posed Problems." *Matematicheskii Sbornik* 61 (103): 211-23.

[4] Lavrentiev, M. M. 1967. *Some Improperly Posed Problems of Mathematical Physics.* Vol. 11. Berlin-Heidelberg: Springer. doi:10.1007/978-3-642-88210-4.

[5] Engl, H. W., Hanke, M., and Neubauer, A. 1996. *Regularization of Inverse Problems.* Dordrecht: Kluwer

Academic Publishers. 321 pp.
[6] Kabanikhin, S. I. 2008. "Definitions and Examples of Inverse and Ill-Posed Problems." *Journal of Inverse and Ill-Posed Problems* 16 (4): 317-57. doi:10.1515/JIIP.2008.019.
[7] Belov, P. A. 2026. "Foundations of the Genetics of Elastoplastic Materials." *Journal of Civil Engineering and Architecture* 20: 129-46. doi:10.17265/1934-7359/2026.04.001.
[8] Belov, P. A., and Golovina, N. Ya. 2026. "Structural Theory of Elastoplastic Material Models." *Journal of Civil Engineering and Architecture* 20: 281-91. doi:10.17265/1934-7359/2026.07.003.

\