Journal of Civil Engineering and Architecture 20 (2026) 347-358
doi: 10.17265/1934-7359/2026.08.004

The Second Inverse Problem of Genetic Identification of Materials: Selection of the Optimal Model Architecture

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: This study formulates the second inverse problem of genetic identification of materials as a natural generalization of the first inverse problem, in which the governing differential operator is reconstructed within a prescribed model architecture. It is shown that selecting the architecture itself constitutes an independent mathematical problem that cannot be reduced to parametric identification. The model space is represented as a discrete set of architectures defined by the smoothness class and the number of segments of the deformation curve. A theorem on the segmentation capacity of an architecture is proved, establishing a class-dependent upper bound on the maximum identifiable number of segments as a function of the experimental sample size. In contrast to the conventional approach, the sample size is treated as a characteristic of the research complex that determines the domain of potential informational observability of architectures. Mathematical admissibility, numerical identifiability, experimental admissibility, and structural optimality of models are consistently distinguished. A functional formulation of the second inverse problem is proposed on the basis of a reduced nonintegrable variational form defined over a continuous representation of the discrete architecture space. Reinsch's principle is generalized: among all experimentally admissible models, preference is given to the architecture of minimum necessary structural complexity. The numerical methodology is based on DOE (design of experiments), interpreted not only as a means of searching for an optimal architecture but also as a tool for designing a research complex capable of providing informational observability of the model space under consideration. The proposed approach is demonstrated using experimental data for SAE 1035 steel.

Key words: Genetic identification, inverse problem, model architecture, smoothness class, design of experiments, operator identification.

1. Introduction

Inverse problems occupy a central place in modern mathematical physics, continuum mechanics, and material identification theory. In their most general form, they seek to reconstruct unknown characteristics of an object from indirect observations or experimental data. The classical formulation of a well-posed mathematical problem was introduced by J. Hadamard, who specified three fundamental requirements: existence of a solution, uniqueness, and continuous dependence on the input data [1]. Violation of at least one of these conditions defines the distinctive character of most inverse problems.

The theory of inverse problems was further developed in the works of A. N. Tikhonov, V. K. Ivanov, M. M. Lavrentiev, and their scientific schools, which established methods for investigating ill-posed problems and laid the foundations of modern regularization theory [2-4]. The concept of an inverse problem has since broadened substantially and now includes the reconstruction of differential-operator coefficients, internal inhomogeneities, sources, domain boundaries, and other characteristics of mathematical models from experimental data [5, 6].

In material mechanics, the most common formulation is the parametric inverse problem. The mathematical model of the material is assumed to have been selected in advance, its structure is regarded as known, and the

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.

experiment is used exclusively to determine a finite set of unknown material parameters. This approach has proved highly effective and underlies most contemporary identification methods. Its domain of applicability, however, is fundamentally restricted to cases in which the selected mathematical model adequately represents the internal architecture of the material under study.

If the model structure is selected incorrectly, no improvement in optimization algorithms can eliminate the resulting structural error. Minimization of a discrepancy functional can identify only the best representative within a model family fixed in advance. Consequently, the attained error simultaneously reflects the quality of optimization, experimental accuracy, and the degree to which the selected model architecture corresponds to the actual material.

This issue provided the starting point for the research program developed by the authors. In "Foundations of the Genetics of Elastoplastic Materials", the concept of materials genetics was introduced, according to which the central object of material theory is an ordinary differential equation for the stress function σ(ε), whereas the object of identification is not an isolated set of material coefficients but the governing ODE operator and the associated system of fundamental solutions [7]. The concept was further developed in "Structural Theory of Elastoplastic Material Models", where different mathematical models were shown to admit a natural interpretation as distinct architectures of governing operators [8].

The next stage of this program was JCEA-3, in which the first inverse problem of genetic identification of materials was formulated. Unlike the classical parametric setting, the reconstructed object was the complete governing differential operator for a prescribed smoothness class and model architecture. The smoothness class was shown to determine uniquely the operator order, the number of primary second-order operators, and the dimension of the fundamental-solution space, whereas the characteristic numbers, factorization, and expansion coefficients must be reconstructed from the experimental stress-strain curve. In addition, a canonical interpolation curve was introduced to provide a unified geometric reference for comparing all models within a fixed smoothness class [9].

At the same time, the results of JCEA-3 revealed a natural limit of applicability of the first inverse problem. The same experiment may admit solutions for different combinations of smoothness class and number of segments, i.e., within different model architectures. The first inverse problem contains no criterion for choosing among such architectures because its formulation assumes that the architecture has already been prescribed. Once the governing operator has been reconstructed, a new and independent problem therefore arises: selection of the model architecture itself. This problem is treated here as the second inverse problem of genetic identification of materials and constitutes the subject of the present study.

In the proposed approach, a model architecture is specified not by a continuous set of parameters but by a discrete position in the space of admissible structures. In the simplest formulation, this space is constructed in the plane "smoothness class-number of segments". Each admissible pair (C^k, s) is associated with a particular model architecture $M(k, s)$, where k denotes the smoothness class and s is the number of segments of the deformation curve. The architecture may therefore be regarded as a discrete function defined on the set of admissible nodes of this plane.

A change in the smoothness class modifies the local structure of the model: the order of the governing differential operator, the dimension of the fundamental-solution space, and the number of conditions required to close the boundary-value problem. A change in the number of segments does not alter the local space of fundamental solutions, but it changes the global representation of the deformation curve because each segment is described by its own linear combination of the same fundamental solutions. Thus, the coordinates k and s characterize two distinct and independent forms

of model structural complexity.

To introduce the genetic terminology used below, consider a governing linear differential operator of even order $2m$. At the first level of factorization, it is represented as the product of m primary second-order operators with real coefficients

$$\mathcal{L}_{2m} = \mathcal{A}_1\mathcal{A}_2\cdots\mathcal{A}_m.$$

The eigenfunctions of these primary second-order operators are termed alleles. This term is introduced not as an independent biological analogy, but as a concise designation for structural components of the fundamental-solution space of the complete governing operator. Each primary second-order operator has the form

$$\mathcal{A}_j = D^2 + a_jD + b_j,$$

where D denotes the differentiation operator, while the coefficients a_j and b_j are real.

At the second level, each primary second-order operator is factorized into two linear operators:

$$\mathcal{A}_j = (D - \lambda_{j1})(D - \lambda_{j2}),$$

where the characteristic numbers λ_{j1} and λ_{j2}, may, in general, be real, complex-conjugate, or repeated. The eigenfunctions of these linear operators are termed genes of the model. Accordingly, the original even-order operator possesses the following two-level operator structure:

$$\begin{aligned} &governing\ operator \\ &\quad \rightarrow primary\ second \\ &\quad - order\ operators \\ &\quad \rightarrow linear\ operators, \end{aligned}$$

which corresponds to the following two-level structure of the generated fundamental solutions:

$$\begin{aligned} &fundamental - solution\ space \rightarrow alleles \\ &\quad \rightarrow genes. \end{aligned}$$

Such a sequential introduction of terminology preserves a rigorous connection between the genetic interpretation and the classical theory of linear differential operators while preventing any confusion between operators and the functions generated by them. Allele operators are the real quadratic factors of the complete governing operator, whereas the alleles themselves are the corresponding eigenfunctions. Gene operators are the linear factors of the allele operators, whereas genes are the fundamental solutions generated by these linear operators.

The first inverse problem formulated in Belov and Weiss [9] is solved independently for each fixed architecture. Its solution reconstructs the governing operator together with its primary and linear factors, characteristic numbers, and expansion coefficients within the selected node of the architecture space. Nevertheless, the same experiment may be represented by several models corresponding to different pairs (C^k, s). Consequently, the first inverse problem does not determine which reconstructed architecture is the most adequate for the investigated material.

Thus, whereas the first inverse problem answers the question, “Which governing operator corresponds to the experiment within a fixed architecture?”, the second inverse problem answers a fundamentally different question: “Which node of the discrete architecture space, i.e., which pair smoothness class-number of segments, provides the most adequate representation of the investigated material?” This transition from optimization of continuous parameters within a single model to comparison of a discrete family of architectures constitutes the central objective of the present paper.

2. Limitations of the First Inverse Problem

In the previous study [9], the first inverse problem of genetic identification of materials was formulated as reconstruction of the complete governing differential operator L within a prescribed architecture, where k denotes the smoothness class and s denotes the number of segments of the model curve. This formulation is fundamentally broader than conventional parametric identification because it reconstructs the operator, its factorization, characteristic numbers, and expansion coefficients over the fundamental solutions. At the same time, fixing the architecture is both a necessary condition of the formulation and its first fundamental limitation.

For a fixed architecture, the smoothness class k determines the operator order, the number of primary second-order operators, and the dimension of the fundamental-solution space. The number of segments s determines the number of local representations of the experimental curve. The internal continuous parameters are optimized only after the pair (k, s) has been selected. Therefore, the first inverse problem defines operations within a single node of the architecture space but contains no rule for transitions between different nodes.

The second limitation is associated with the research complex. In the present study, the research complex is understood as an integrated experimental-computational system that combines the measurement system and loading program, the number and distribution of experimental points, instrumental accuracy, and the computational procedures used to reconstruct and test admissible model architectures. It therefore characterizes not the material itself, but the amount, accuracy, and structure of experimentally distinguishable information available for identification. An architecture may be mathematically well defined but remain unidentifiable from the available sample. Consequently, reconstruction of the governing operator is constrained not only by the selected architecture but also by the informational capacity of the research complex.

Here, the experimental program specifies the prescribed loading history and the set of measured physical quantities. The loading program defines the sequence, magnitude, and direction of the imposed mechanical actions, whereas the measurement channels specify which physical responses are recorded simultaneously during this loading. Thus, the experimental program, together with the number and distribution of sampling points and the instrumental accuracy of the channels, determines the experimentally available information supplied to the identification procedure.

Let N denote the number of experimental observations. It should be regarded not merely as the number of rows in a data table but as a characteristic of the research complex itself. If N is insufficient, the corresponding architecture is fundamentally unobservable regardless of the optimization procedure. Even when N is sufficient, issues of rank, conditioning, and stability of identification remain.

Accordingly, the results of the first inverse problem are conditional: they are associated simultaneously with a fixed architecture and a fixed research complex. Transition to architecture selection therefore requires construction not of all formally conceivable models but only of the set of mathematically admissible and informationally observable architectures.

3. Discrete Space of Admissible Architectures and Its Continuous Representation

The discrete architecture space is defined as the set of models $M(k,s)$, where $k = 0,1,2,\ldots$ is the smoothness-class index C^k, and $s = 1,2,\ldots$ is the number of segments of the model curve. Each admissible pair (k, s) specifies an individual architecture within which an independent first inverse problem is solved.

Changing k modifies the local operator structure: the order of the governing operator L, the number of primary second-order operators, the number of linear factors, and the dimension of the fundamental-solution space. Changing s does not alter the local fundamental-solution space but changes the global number of intervals, on each of which the experimental curve is represented by its own linear combination of these solutions.

Thus, k and s characterize two different types of structural complexity—operator complexity and segmentation complexity. However, they are not independent admissibility coordinates: increasing the smoothness class increases the number of conditions required to determine each segment and thereby reduces the maximum admissible segmentation for a fixed sample size.

3.1 Theorem 1. Minimum Sample Size and Segmentation Capacity

Let the smoothness class C^k correspond to a

governing differential operator of order $2m$, where $m = k + 1$.

Each contact node of adjacent spline segments of class C^k requires $k + 1$ independent informative experimental points located within the corresponding segment. These points provide the information required to satisfy the continuity conditions for the function and its derivatives through order k.

Consequently, each segment must provide $k + 1$ informative points for each of its two nodes, giving the minimum informational capacity

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

Since adjacent segments share a common contact node, the minimum sample size required to identify an s-segment spline is

$$N_{min} = (s + 1)(k + 1) \quad (1)$$

Therefore, the necessary condition for potential identifiability is

$$N \geq (s + 1)(k + 1) \quad (2)$$

For a prescribed sample size N, the maximum admissible number of segments is

$$s_{max} = \left\lfloor \frac{N}{k + 1} \right\rfloor - 1 \quad (3)$$

Conversely, if the smoothness class C^k and the required maximum number of segments s_{max} are specified in advance, the minimum sample size for experimental planning is

$$N_{min} = (s_{max} + 1)(k + 1) \quad (4)$$

For example, a six-segment spline of class C^2 requires

$$N_{min} = (6 + 1)(2 + 1) = 21$$

Thus, the experimental sample must contain at least 21 independent informative points. Conditions (2) and (4) are necessary information-balance conditions for potential identifiability; they are not, by themselves, conditions of numerical identifiability. They determine whether the experimental sample contains a sufficient number of independent informative points for the prescribed architecture. Numerical identifiability is a stronger requirement and must additionally be verified from the actual identification system by its rank, conditioning, and stability with respect to admissible perturbations of the experimental data. In this sense, Conditions (2) and (4) precede, rather than replace, the numerical-identifiability criterion introduced for the first inverse problem in Belov and Weiss [9].

The model remains denoted by M(k, s), because N is not a property of the material or of the architecture. The sample size determines the domain of architectures accessible to a particular research complex:

$$D\_N = \{(k, s): k \in K, 1 \leq s \leq s\_max(N, k)\} \quad (5)$$

The set D_N is the domain of potential informational observability of architectures. Increasing N expands this domain; decreasing N or thinning the data contracts it.

4. The Second Inverse Problem of Genetic Identification of Materials

The first inverse problem answers the question: which governing operator corresponds to the experiment within a fixed architecture? The second inverse problem answers a different question: which architecture from the set accessible to the research complex should be selected to represent the investigated material?

The second inverse problem consists in constructing the admissible family of architectures, solving the first inverse problem successively at the admissible nodes, and selecting the model that is mathematically admissible, numerically identifiable, experimentally admissible, and minimally sufficient in structural complexity.

These requirements form the following hierarchy:

$$\begin{aligned} &mathematical\ admissibility \\ &\rightarrow numerical\ identifiability \\ &\rightarrow experimental\ admissibility \\ &\rightarrow structural\ optimality \end{aligned} \quad (6)$$

An architecture is mathematically admissible for a sample of size N if the segmentation-capacity condition is satisfied.

$$s \leq s\_max(N, k) \quad (7)$$

Condition (7) is verified before launching the continuous optimizer. Architectures outside the admissible domain are excluded from the computational plan because the amount of independent information is

already insufficient at the level of the necessary information balance.

Mathematical admissibility alone does not guarantee stable reconstruction. For every candidate architecture, the local identification system must be examined for full rank, acceptable conditioning, and stability with respect to small perturbations of the experimental data. Preprocessing maps the experimental stress-strain sample to the unit square, so both strain and stress are dimensionless and normalized. Consequently, the instrumental error Δ directly sets the perturbation scale of the normalized data. If A denotes the scaled local identification matrix, its conditioning is regarded as acceptable when the amplification of perturbations compatible with the instrumental accuracy remains subcritical,

$$\kappa(A)\Delta < 1$$

so that the natural conditioning limit is determined by the experimental accuracy itself, $\kappa(A) < \frac{1}{\Delta}$, rather than by an independently prescribed numerical threshold. The product $\kappa(A)\Delta$ is therefore used as a continuous indicator of identification stability: the smaller it is relative to unity, the weaker the amplification of instrumental perturbations.

Local subdivision is permitted only when the information-balance condition for the corresponding smoothness class is satisfied, the local system has full rank, and its conditioning remains acceptable. A reduction of the objective function at the experimental points is insufficient; the daughter spline is additionally verified on a dense control grid to exclude hidden oscillations between neighboring measurement points.

Every candidate architecture is associated with a reconstructed stress-strain curve. A model is regarded as experimentally admissible if the entire reconstructed curve lies within the instrumental-error tube.

$$\left|\sigma_{mod}(\varepsilon_i; k, s) - \sigma_{exp}(\varepsilon_i)\right| \le \Delta\sigma_i, \quad i = 1, \dots, N \tag{8}$$

For each fixed smoothness class, the search begins with a single segment. If Criterion (8) is violated, one segment is greedily subdivided and the first inverse problem is solved again only for the new architecture while all unchanged segments are inherited.

For each fixed smoothness class, let s_{opt} denote the first experimentally admissible number of segments, i.e., the first value at which mathematical admissibility, numerical identifiability, and complete immersion into the instrumental-error tube are simultaneously achieved. The search is bounded by the informational segmentation-capacity limit $s_{max}(N, k)$, so that $s_{opt} \le s_{max}(N, k)$. Equality occurs only when the first experimentally admissible architecture is reached exactly at the information-capacity boundary.

If experimental admissibility cannot be achieved at the segmentation-capacity limit, further subdivision is impossible for the given research complex. This situation indicates exhaustion of the sample segmentation capacity but does not by itself prove the physical inadequacy of the smoothness class.

The residual discrepancy may result either from insufficient approximation capability of the smoothness class or from insufficient informational capacity of the experiment for testing more segmented architectures. Distinguishing between these causes requires increasing the sample size, changing the point distribution, or introducing additional measurement channels.

Therefore, transition to the next smoothness class must be accompanied by an explicit diagnosis: “the class is insufficient for the given research complex”. An absolute conclusion regarding class inadequacy is justified only after verifying the stability of this conclusion with respect to enrichment of the experimental information.

For every sufficient smoothness class, the algorithm determines s_{opt}, the first experimentally admissible number of segments. Several classes may yield models located inside the same instrumental-error tube; therefore, experimental admissibility is necessary but not sufficient for optimality.

Increasing k raises operator complexity, whereas increasing s raises segmentation complexity. The optimal model is not necessarily characterized by the minimum k, minimum s, or RMSE (minimum root-mean-square error). The second inverse problem selects the best balance between these two kinds of structural complexity while satisfying Conditions (7) and (8) together with the requirements of numerical identifiability.

5. Lower Bound and the Generalized Reinsch Principle

For a fixed architecture, a fixed experimental sample, an error metric, and an admissible space of internal parameters, the minimum attainable error is defined as the exact lower bound over all admissible values of the internal parameters:

$$RMSE_{min}(k, s \mid N, E) = inf_{\theta \in \Theta\_adm(k,s;E)} RMSE(k, s; \theta \mid N, E) \quad (9)$$

where θ denotes the vector of continuous parameters of the first inverse problem, $\Theta\_adm(k, s; E)$ is its admissible parameter set for the prescribed architecture and research complex, and E is the complete set of characteristics of the research complex, including the distribution of measurement points and the experimental program. This quantity characterizes the approximation capability of the architecture with respect to the particular sample, but it is not an absolute property of the material.

5.1 Theorem 2

For fixed experimental data, a fixed research complex, a fixed metric, a prescribed architecture, and the same admissible space of internal parameters, the global value of the lower bound is independent of the optimization algorithm used to attain it.

Proof. The architecture defines the same set of representable curves. Different algorithms may exhibit different convergence rates and may terminate at different local minima; however, upon reaching the global minimum, they minimize the same function over the same admissible set. Therefore, the exact lower bound is unique. A change in N, in the distribution of points, in the metric, or in the admissible parameter set constitutes a change in the problem formulation and does not contradict the theorem.

The lower bound diagnoses the limit of the architecture's approximation capability for the given sample. The instrumental error defines an external tube of experimental distinguishability. These quantities are not interchangeable: the former depends on the model and the data, whereas the latter is determined by the measurement system.

If the globally optimized model leaves the error tube, the architecture is insufficient for the given research complex. When the first experimentally admissible architecture is reached, the corresponding number of segments is s_{opt} for the prescribed smoothness class. A more complex architecture may have a smaller error, but this advantage may be experimentally indistinguishable.

The generalized Reinsch principle is formulated as follows: among mathematically admissible and numerically identifiable models that agree with the experiment within the instrumental error, the optimal model is the architecture of minimum necessary structural complexity.

Experimental admissibility is a constraint, whereas minimum structural complexity is the optimality criterion. Once the model has entered the error tube, a further reduction in error does not by itself justify a more complex architecture. The admissible search region is restricted in advance by the informational capacity of the research complex.

6. Functional Formulation of the Second Inverse Problem

For a fixed architecture, the first inverse problem reconstructs the stress and the internal parameters from the stationarity condition of the original nonintegrable variational form. The resulting solution is substituted into the original form, after which internal integration

with respect to strain is carried out. The strain variable is eliminated, while the variational nature of the formulation is retained.

This procedure yields a reduced nonintegrable variational form. We denote this reduced form by

$$\delta\Phi \equiv \delta\Phi(k, s; E), \quad (10)$$

where E denotes the fixed research complex. Form (10) inherits the reversible part, the stress derivatives, and the dissipation channels of the original formulation; it is not an ordinary function and cannot be reduced to the squared deviation of the model curve from the experiment.

Since k and s are discrete integer coordinates of the architecture space, infinitesimal variations with respect to them are not defined on the original discrete set. To formulate the second inverse problem, the values of the reduced form at admissible nodes are extended to a continuous surface over the admissible region. Extension outside this region has no informational justification.

Points with noninteger coordinates are auxiliary continuation points rather than independent physical models. They make it possible to estimate the local directions of variation of the form and the induced stress variations with respect to the architecture parameters.

$$\delta\sigma = (\partial\sigma/\partial k)\delta k + (\partial\sigma/\partial s)\delta s. \quad (11)$$

After substitution of the induced variations, the reduced form can be represented as

$$\delta\Phi = \Phi_k\,\delta k + \Phi_s\,\delta s. \quad (12)$$

Here Φ_k and Φ_s denote the coefficients of the independent architecture variations δk and δs in the continuous representation of the reduced variational form.

Because the variations δk and δs are independent, a stationary architecture of the continuous representation satisfies the conditions

$$\Phi_k = 0, \ \Phi_s = 0. \quad (13)$$

The stationary point may have noninteger coordinates. Therefore, the final solution is selected among the nearest real architectures with integers k and s belonging to the admissible domain. For every candidate, the first inverse problem is solved again and the rank, conditioning, pointwise admissibility, and structural complexity are verified.

The sample size N, instrumental error Δ, and point-placement scheme are not introduced as variational coordinates of the model. They are parameters of the research complex E and determine the admissible domain, the form of the available data, and the criterion of experimental admissibility. Consequently, changing E creates a new representation of the architecture space and may change the selected optimal model without changing the material itself.

This distinction is fundamental: material properties are invariant with respect to the measurement system, whereas the observable and optimal architecture of the mathematical model depends on the amount of experimentally distinguishable information.

7. DOE as a Numerical Realization of the Second Inverse Problem

DOE (design of experiments) is used as a mechanism for controlling the search in architecture space. Within each fixed architecture, the continuous optimizer of the first inverse problem is employed. In view of the role of the research complex, two levels of experimental design should be distinguished. The term levels is essential here: these are two hierarchically related formulations of the design problem, not two mandatory sequential stages of a single computational procedure.

At the first level—architecture DOE—the factors are k and s, and the design is restricted to the admissible domain. At the second level—DOE of the research complex—the factors may include the sample size N, the distribution of points, the instrumental error Δ, and the loading program. The first level selects an architecture within the information supplied by a prescribed research complex; the second level designs or modifies the research complex itself and thereby determines which part of the architecture space can be observed. Thus, the two levels are logically coupled but

need not be executed as successive stages in every application.

In the present study, the principal computational cycle belongs to the first level. The second level is formulated as a design principle and is investigated through the analysis of segmentation capacity and minimum sample size.

For a prescribed research complex, with K denoting the set of admissible smoothness-class indices, only the following set of nodes is generated:

$$D_N = \{(k,s): k \in K, 1 \le s \le s_{max}(N,k)\}. \quad (14)$$

Each computational run includes selection of an architecture, verification of the information balance, solution of the first inverse problem, calculation of the error measure, verification of rank and conditioning, shape control on a dense grid, and verification that the curve lies within the instrumental-error tube. Continuous operator parameters are not external DOE factors; they are optimized within each run.

For each smoothness class, the computation begins with a single segment. At every iteration, one parent segment is selected whose replacement by two daughter segments yields the greatest local improvement subject to the admissibility conditions. All other segments are inherited without recomputation.

For each candidate subdivision, the scaled local identification system associated with the two daughter segments is examined before the subdivision is accepted. A subdivision candidate is rejected if the information balance is violated, the local system loses full rank, the conditioning criterion $\kappa(A)\Delta < 1$ is violated, or the control grid reveals oscillations hidden between sample points. Thus, increasing the number of segments by one is a necessary consequence of the operation, but not proof of its correctness.

The DOE responses are the error measure, the maximum pointwise deviation, the indicator of immersion into the error tube, the rank and condition number of the local systems, the measures of operator and segmentation complexity, and the coefficients of the reduced variational form. A continuous representation is constructed from the computed nodes only inside the admissible domain.

Stationary regions are located by detecting zeros or sign changes of the corresponding coefficients. The resulting continuous points serve as centers for a local search among the nearest integer architectures, after which the generalized Reinsch principle is applied.

When planning a new experiment, Eq. (4) makes it possible to determine in advance the minimum sample size for the architecture range under consideration. A practical design, however, should include an information margin above this minimum because the necessary balance does not guarantee stability. The responses of the second DOE level include the size of the admissible domain, the maximum accessible segmentation, conditioning, robustness of the selected architecture under data thinning, and the probability that the optimal node changes when the measurement system is modified.

Such a DOE answers not only the question “Which model should be selected?” but also “Which research complex is required to distinguish between competing architectures?” Experimental design and model selection therefore form a coupled problem.

For demonstration, the range k = 0, 1, 2 is considered. For each class, the model family is generated only until experimental admissibility is first achieved or the segmentation-capacity boundary is reached. The highest class considered here defines the upper limit of the present computational demonstration. In a future study, the smoothness class will be fixed and the foundations for classification of elastoplastic materials will be developed using a canonical six-segment spline template of the stress-strain curve.

The numerical demonstration uses an experimental sample for SAE 1035 steel. The same absolute instrumental error $\Delta = 0.01$ is used for all three smoothness classes. Within each class, the first admissible architecture is the first one satisfying the pointwise Condition (8).

Table 1 Informational segmentation-capacity limit and first experimentally admissible number of segments for classes C^0-C^2 for SAE 1035 steel at Δ = 0.01.

Class	s_{max}(N, k)	s_{opt}	RMSE	max |Δσ|	Error tube
C^0	176	16	0.002296	0.009333	Yes
C^1	87	4	0.003272	0.009269	Yes
C^2	58	4	0.002042	0.008058	Yes

For the present SAE 1035 sample, N = 177. According to Eq. (3), the theoretical segmentation-capacity limits are

$$s_{max}(177,0) = 176,\ s_{max}(177,1) = 87,\ s_{max}(177,2) = 58 \quad (15)$$

Table 1 combines analytical information-capacity bounds with the results of the architecture-search DOE. The values $s_{max}(N, k)$ are calculated directly from Eq. (3) for the prescribed sample size and are not DOE responses. In contrast, s_{opt}, RMSE, the maximum pointwise deviation, and the error-tube indicator are obtained from the sequential computational exploration of admissible architectures, which constitutes the first-level architecture DOE in the present formulation.

The first experimentally admissible architectures have s_{opt} = 16, 4, and 4 for classes C^0, C^1, and C^2, respectively, whereas the corresponding informational limits $s_{max}(N, k)$ are 176, 87, and 58. Thus, all three selected architectures lie well inside their admissible domains and satisfy $s_{opt} < s_{max}(N, k)$ for the present sample.

The minimum sample sizes required by Eq. (4) for potential identification of the resulting architectures are 17, 10, and 15, respectively. These values are necessary informational bounds. They do not guarantee stable reconstruction and do not replace verification of rank, conditioning, and point distribution.

Raising the smoothness class reduced the first experimentally admissible number of segments s_{opt} from 16 to 4. The subsequent transition to the next class did not reduce s_{opt} at the selected error level, although it decreased the maximum pointwise deviation. Because both models lie inside the same error tube, the smaller error alone does not justify selecting the more complex class; the final decision must account for operator complexity in accordance with the generalized Reinsch principle.

The example also shows that the available sample size N = 177 substantially exceeds the minimum informational bounds for the identified architectures: 17 points for C^0 with 16 segments, 10 points for C^1 with 4 segments, and 15 points for C^2 with 4 segments. SAE 1035 steel is therefore suitable not only for demonstrating model selection but also for a subsequent data-thinning DOE, in which the stability of the optimal architecture can be investigated as a function of sample size and point distribution.

8. Conclusion

The second inverse problem of genetic identification of materials has been formulated. It arises from a fundamental limitation of the first inverse problem: the governing operator can be reconstructed only within a preselected architecture, whereas the architecture itself remains an external condition of the formulation.

The model space is shown to consist of a discrete set of nodes in the plane “smoothness class-number of segments”. For every fixed smoothness class, the greedy refinement algorithm constructs a family of models ranging from a single segment to the first model completely contained within the instrumental-error tube.

The criterion for the optimal number of segments is not a prescribed error value but the first attainment of experimental admissibility, which defines s_{opt} for the prescribed smoothness class. The informational limit $s_{max}(N, k)$ bounds the search from above. Reaching this segmentation-capacity limit without entering the error tube indicates insufficiency of the selected smoothness class for the given research complex.

The minimum attainable RMSE is interpreted as an invariant of the architecture with respect to the

optimization algorithm for fixed data, metric, and admissible internal-parameter space. The error is thereby transformed from an ordinary numerical measure into a diagnostic characteristic of the model architecture.

The generalized Reinsch principle is adopted as the mathematical basis for selection: among models consistent with the experiment within the instrumental error, preference is given to the architecture of minimum necessary structural complexity.

The first inverse problem reconstructs the operator L, its factorization, and its fundamental solutions within a fixed architecture node. The second inverse problem investigates transitions and operations between nodes and determines the optimal architecture.

Substitution of the solution of the first inverse problem into the original nonintegrable variational form, followed by internal integration with respect to strain, does not convert it into an ordinary functional. Instead, it produces a reduced nonintegrable variational form defined on the continuous surface of models.

Continuous extension of the discrete space is required to define independent variations with respect to k and s and to formulate the stationarity conditions of the second inverse problem. The final selection is nevertheless made only among real models with integer coordinates.

DOE is interpreted as a numerical mechanism for constructing a discrete sample of architectures, solving the first inverse problem at the nodes, reconstructing the continuous surface, and locating stationary models. A particular realization has been considered for the selected range of smoothness classes.

The first inverse problem reconstructs the real coefficients of the allele operators, in which the reversible modules A_{ij} and dissipative modules B_{ij} enter as inseparable combinations. Their separate determination even from a single monotonic stress-strain curve is fundamentally impossible.

Separation of the reversible and dissipative modules requires additional thermodynamic cycles in which loss coefficients are determined from the oriented areas enclosed by closed trajectories in phase space. This formulation constitutes the third inverse problem and lies beyond the scope of the present study.

A future study will consider a special case of the general theory at a fixed smoothness class and will establish the foundations for classification of elastoplastic materials using a canonical six-segment spline template of the stress-strain curve. In that formulation, the focus will shift from architecture optimization to material classification and recognition within the prescribed six-segment architecture.

Thus, material properties are invariant with respect to the choice of measurement system, whereas the optimal architecture of the mathematical model is determined not only by the smoothness class and the number of segments but also by the instrumental error of the measurement system, which determines the amount of experimentally distinguishable information.

The sequence of studies forms a unified program: the first inverse problem reconstructs the operator within a fixed model; the second inverse problem determines the optimal architecture; and the subsequent recognition problem identifies the material in the space of preselected architecture templates.

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.

[9] Belov, P. A., and Weiss, E. 2026. "Direct and First Inverse Problems of Genetic Identification of Materials at a Fixed Smoothness Class." *Journal of Civil Engineering and Architecture.* (Accepted for publication).