Interpolation and Optimization through Gagliardo Diagrams

Björn Jawerth Mario Milman*

Abstract

We develop an interpolation method for solving constrained optimization problems on admissible sets with no linear structure. The starting point is the Gagliardo diagram of attainable cost and error pairs. When the two criteria are sums of component contributions, geometric selection on the component diagrams produces a weighted approximation problem. The weights record successive cost increases and error reductions, and reconstruction maps turn solutions of the weighted problem into admissible choices for the original one. Separate control of cost and error gives a constructive, variational form of the fundamental lemma of interpolation, applicable simultaneously to all constraint levels.

For normalized finite data, the optimal error functions F of the original problem and D of the weighted model satisfy

$$D(B) \le F(B) \le 2D(B/2), \qquad B \ge 0.$$

Thus the representation provides both lower bounds for the optimal error and admissible approximations with explicit upper bounds. No convexity is required. Under closedness and summability assumptions, the construction extends to countable families, with uniform control of the error from finite truncations. The knapsack problem is an exact instance of the correspondence.

The same geometric viewpoint also clarifies the relation between constrained and penalized optimization. Lower convex hulls represent penalized minima exactly by weighted K-functionals and give nested families of approximate solutions. Power penalties recover the constrained information that linear penalties may lose.

Keywords. Knapsack; interpolation; Gagliardo diagrams; sharp constants; maximum error; countable representations; nonlinear approximation.

2020 Mathematics Subject Classification. Primary 46B70; Secondary 41A65, 90C27, 90C29.

*Björn Jawerth passed away in 2013. This paper publishes part of the material prepared jointly for the Paseky lectures and intended for subsequent publications. The lecture manuscript was never published; the Charles University booklet contained only its introduction. Completion of this work was considerably delayed. The second author takes responsibility for the preparation of the final version and for any errors or shortcomings.

1 Introduction

An optimization problem often involves a tension between two competing requirements: keeping the cost small and making the error small. A more accurate choice may require greater resources, while a cheaper choice may leave a larger error. Interpolation theory begins with a closely related problem. An element is decomposed into two parts, whose sizes are measured in different spaces, and one seeks a balance between these sizes. In both settings, the attainable pairs describe how small one quantity can be made when the other is constrained.

Micchelli and Pinkus [10] developed the use of Gagliardo diagrams in optimization to study competing variational criteria. Their work showed how the geometry of attainable pairs relates constrained and penalized formulations and helps describe their solutions. This viewpoint is a starting point for the present paper.

Building on this viewpoint, we develop a constructive interpolation method for constrained optimization. The method applies to sums of cost and error criteria on component sets that need no linear or convex structure. Its basis is the geometry of attainable cost and error pairs: from these pairs we construct a weighted approximation problem and maps that transfer choices between the two problems. The maps control cost and error separately, so a solution of the weighted problem produces an admissible choice for the original problem with explicit bounds. One construction works for every constraint level.

The object of the method is the optimal error as a function of the available resources. Representing this function makes it possible to compare allocations across all budgets, transfer approximate minimizers, and relate constrained and penalized optimization. This is useful when a common resource must be distributed among many choices: coordinates of an approximant, accuracy levels in component problems, or corrections from a hierarchical decomposition. The original choices may be discrete or continuously parametrized; the representation is built from their costs and errors.

Gagliardo's diagrams provide the geometric setting [6, 7]. For nonnegative functionals C, E on an admissible set $\mathcal{A}$, the upward diagram is

$$\Gamma = \{(s,t) \in [0,\infty)^2 : C(z) \le s,\ E(z) \le t \text{ for some } z \in \mathcal{A}\}.$$

Its lower boundary describes the optimal error under a cost constraint, while its supporting lines determine the minima of $C + \lambda E$. Micchelli and Pinkus [10, Section 2] analyze this relationship under convexity assumptions. Our construction associates weighted interpolation models with the component diagrams and recovers original admissible choices with separate bounds for cost and error.

The 0/1 knapsack problem gives an exact instance of the method. For positive weights c_j and p_j, consider

$$V(B) = \max\left\{\sum_{j=1}^{J} p_j\xi_j : \sum_{j=1}^{J} c_j\xi_j \le B, \quad \xi \in \{0,1\}^J\right\}. \tag{1.1}$$

Writing $P_* = \sum_j p_j$, the corresponding approximation error is

$$P_* - V(B) = \min_{\substack{\xi\in\{0,1\}^J \\ \sum_j c_j\xi_j \le B}} \sum_j p_j(1-\xi_j) = E(B, \mathbf{1}; \ell^0(c), \ell^1(p)). \tag{1.2}$$

Here $\|z\|_{\ell^0(c)} = \sum_{j:z_j\neq 0} c_j$ is the weighted support size and $E(B, f; U, V) = \inf_{\|z\|_U \le B} \|f - z\|_V$. Replacing every nonzero coordinate of an approximant to $\mathbf{1}$ by one preserves its support and does

not increase its error. Thus one fixed weighted pair represents the knapsack problem exactly for every budget. The general theorem constructs a model of this kind for arbitrary normalized finite component data, with uniform comparison factors and maps returning model solutions to original admissible choices.

The main construction selects points on each component Gagliardo diagram so that costs increase and errors decrease geometrically. Successive differences become the costs and gains of a weighted binary problem. An arbitrary selection of these increments need not describe an original admissible choice. The decisive step is a replacement principle: the selection is converted into an initial segment in each component, with a controlled increase in cost and no increase in model error. Each initial segment determines a component level and hence an admissible choice. This is the reconstruction mechanism behind the representation, and the sharp cost of this replacement is one of the questions settled here.

The strong fundamental lemma of Cwikel, Jawerth, and Milman [3] provides the guiding principle. Under the Banach-pair hypotheses recalled in Section 10.3, it gives a single representation

$$f = \int_0^\infty U(s)\,\frac{ds}{s}$$

for which

$$K(\lambda, f; \mathbf{X}) \le \int_0^\infty \min\{1, \lambda/s\} J_{\mathbf{X}}(s, U(s))\,\frac{ds}{s} \le cK(\lambda, f; \mathbf{X}), \qquad \lambda > 0,$$

where $\mathbf{X} = (X_0, X_1)$ and $J_{\mathbf{X}}(s,u) = \max\{\|u\|_{X_0}, s\|u\|_{X_1}\}$. The same U works for every λ. Splitting its integral at different scales supplies a whole family of decompositions with controlled endpoint sizes. This simultaneous control of the entire K-functional is the feature of the fundamental lemma that we bring to optimization.

Our variational construction develops this principle directly from the component diagrams. A two-parameter version of Gagliardo's geometric selection gives scalar cost increases and error reductions, which become the weights of the approximation model. The role of summing representing vectors is played here by reconstruction: selected increments are replaced by initial segments and converted into original admissible choices. This step requires no addition or subtraction on the original admissible sets. Separate estimates for cost and error pass to sums of components, with constants independent of the number of components and choices. Thus the same representation and the same comparison maps work for every constraint bound and penalty parameter. Section 10.3 also gives a direct Banach-pair realization of the connection: the lower convex hull weights represent the penalized minima by a weighted K-functional, and logarithmic averaging gives a continuous K–J representation.

The work of Domínguez and Milman [4] provides further support for this program of applying interpolation methods to optimization and variational problems. Their use of variational K-functionals, Gagliardo relative completions, and extrapolation connects Γ-convergence with quantitative estimates for minimizers. The companion manuscript [11] develops another part of this connection. For a class of closed convex homogeneous energies on a Hilbert space, it selects series representations whose terms converge in ℓ^1 under Mosco convergence of the energies and strong convergence of the data. Their prescribed cost functions also converge uniformly over all positive parameters.

These developments suggest studying convergence directly in the more general setting of the present paper. When component costs, errors, or admissible choices vary, one can ask how their attainable diagrams and weighted models converge, and which assumptions ensure convergence of

the optimal error functions. The reconstruction maps add a further question: can model solutions be chosen so that the corresponding admissible approximations converge? Such questions call for a suitable notion of convergence on the varying admissible sets, together with control of their cost–error diagrams. The geometric representations and separate bounds for the two criteria provide a starting point for this investigation; the Γ- and Mosco-convergence results indicate the kinds of stability one can seek.

For the allocation problems studied here, the representation also gives a route from infinite choices to finite optimization. Under closedness and summability assumptions, the construction extends to countably many components and levels. The weights of omitted increments bound the truncation error uniformly over all budgets, so finite model solutions give admissible approximations with explicit control of the omitted contributions.

A complementary construction explains the role of penalization. The lower convex hulls of the component diagrams represent linear penalty minima exactly by weighted K-functionals and produce an ordered family of admissible approximate solutions. Supporting lines can omit choices that are optimal at particular constraint bounds. Power penalties recover this missing information about the constrained error function. The diagram viewpoint thus identifies the information retained by each representation and provides a way to pass between the corresponding optimization problems.

1.1 A variational fundamental lemma

Let $\mathcal{A}_i$, $1 \le i \le r$, be nonempty finite sets, with $C_i, E_i \colon \mathcal{A}_i \to [0,\infty)$ normalized by

$$\min_{\mathcal{A}_i} C_i = \min_{\mathcal{A}_i} E_i = 0.$$

The minima may occur at different points. For $z \in \mathcal{A} = \prod_i \mathcal{A}_i$, put

$$C(z) = \sum_i C_i(z_i), \qquad E(z) = \sum_i E_i(z_i), \qquad F(B) = \min_{C(z)\le B} E(z). \tag{1.3}$$

Subtracting the component minima reduces arbitrary nonnegative finite data to this normalization. Section 6.1 gives the corresponding estimates for the original functionals.

Theorem 1.1 (Variational fundamental lemma)**.** *For every $a, b \ge 2$, there are a finite set J, positive weights A_j, W_j, and maps*

$$P \colon \mathcal{A} \longrightarrow \{0,1\}^J, \qquad L \colon \{0,1\}^J \longrightarrow \mathcal{A}$$

with the following properties. Writing

$$C_d(\xi) = \sum_{j\in J} A_j \xi_j, \qquad E_d(\xi) = \sum_{j\in J} W_j (1-\xi_j),$$

for every $z \in \mathcal{A}$ and $\xi \in \{0,1\}^J$ one has

$$C_d(Pz) \le C(z), \qquad E_d(Pz) \le E(z), \tag{1.4}$$

$$C(L\xi) \le a C_d(\xi), \qquad E(L\xi) \le b E_d(\xi). \tag{1.5}$$

Consequently, with

$$D(B) = \min_{C_d(\xi)\le B} E_d(\xi) = E(B, \mathbf{1}; \ell^0(A), \ell^1(W)),$$

the optimal error function satisfies

$$\boxed{D(B) \le F(B) \le bD(B/a), \qquad B \ge 0.} \tag{1.6}$$

The constants are independent of the number of components and the cardinalities of the admissible sets.

The interpolation structure in Theorem 1.1 is constructed from the optimization data. Binary vectors, with addition modulo two and weights A, W, form the pair of weighted normed Abelian groups used in the model. For real vectors, the corresponding functionals are the weighted support size and the weighted approximation error. The original choices enter through their attainable cost and error pairs and the reconstruction map; no linear operations on these choices are used.

The maps P and L compare individual choices. If a model vector has cost at most B/a and error at most η, its image under L has cost at most B and error at most $b\eta$. This gives an admissible approximation with explicit bounds whenever an exact or approximate model solution is available. Replacing the model support by initial segments and choosing component minimizers defines the reconstruction map.

Adding the component estimates gives the same bounds for the total cost and error. For upward Gagliardo diagrams, this yields

$$\operatorname{diag}(a, b)\, \Gamma_d \subseteq \Gamma \subseteq \Gamma_d.$$

For the penalized values

$$\mathcal{K}(\lambda) = \min_z \{C(z) + \lambda E(z)\}, \qquad \mathcal{K}_d(\lambda) = \sum_j \min(A_j, \lambda W_j), \tag{1.7}$$

they give

$$\mathcal{K}_d(\lambda) \le \mathcal{K}(\lambda) \le a\mathcal{K}_d((b/a)\lambda). \tag{1.8}$$

The same representation therefore controls the constrained values, the penalized values, and the real interpolation quantities formed from them. For example, the $L^q(d\lambda/\lambda)$ quantities associated with $\lambda^{-\theta}\mathcal{K}(\lambda)$ are compared with factor $a^{1-\theta}b^{\theta}$, for $0 < \theta < 1$ and $0 < q \le \infty$; Corollary 6.3 also gives the corresponding approximation estimates. Taking $a = b = 2$ yields factor two in (1.6) and (1.8). For fixed B, computing $D(B)$ is a binary knapsack problem, and L reconstructs an admissible element from each solution.

1.2 Sharp bounds for replacement by initial segments

A component level corresponds to taking all increments up to that level. An arbitrary binary vector may omit an earlier increment and select a later one, so it need not describe an original admissible choice. We determine the least cost increase needed to replace such selections by initial segments without increasing their model error. This identifies the exact constant in a specific step of the reconstruction.

For geometrically selected row nodes (t_n, u_n) with $t_0 = 0$ and

$$t_{n+1} \ge at_n, \qquad u_{n+1} \le u_n/b,$$

put $A_n = t_n - t_{n-1}$ and $W_n = u_n - u_{n+1}$. Let D_{sel} and D_{pre} denote the optimal error functions under the same cost constraint, for arbitrary binary vectors and for vectors supported on an initial segment in each row, respectively. For $a > 1$ and $b \ge 1$, assuming that the error coordinates tend

to zero in every infinite row when $b = 1$, Theorem 5.2 determines the least uniform cost factor for replacement without increasing the error:

$$\rho(a,b) = \begin{cases} a/(a-1), & 1 \le b < 2, \\ \max\{1, 1/(a-1)\}, & b \ge 2. \end{cases} \tag{1.9}$$

Including every coordinate up to the last selected one gives $a/(a-1)$ in Proposition 5.1. For $b \ge 2$, each gain is at least the sum of the later gains, which allows a replacement determined by the first omitted coordinate. Two-item rows show that this uniform factor is sharp for the stated geometric class, under the requirement that error does not increase. The corresponding optimal error functions satisfy

$$D_{\mathrm{sel}}(B) \le D_{\mathrm{pre}}(B) \le D_{\mathrm{sel}}(B/\rho(a,b)). \tag{1.10}$$

For $a, b > 1$, composing this replacement with the map from initial segments to original admissible choices gives

$$D_{\mathrm{sel}}(B) \le F(B) \le bD_{\mathrm{sel}}\Big(\frac{B}{a\rho(a,b)}\Big); \tag{1.11}$$

the cost and error factors for the original choices are $(a\rho(a,b), b)$ (Corollary 7.7). They reduce to (a,b) for $a, b \ge 2$, and to $(a/(a-1), b)$ for $1 < a < 2$, $b \ge 2$. Section 5 develops the replacement arguments of the unpublished Paseky manuscript [9, Section 4.4] and determines the exact factor above. For $a = 3/2$ and $b \ge 2$, the replacement factor is two, improving the integer-scale estimate $9/4$ recalled in Remark 5.4. After reconstruction, the cost factor for the original problem is three.

1.3 Countable families and finite approximations

Approximation problems may offer infinitely many refinement levels in each component, or involve countably many components under a common constraint. In this setting the representation must also describe the error incurred when only finitely many model coordinates are retained. Closedness of the component diagrams and a summable bound for the initial errors give a countable representation with uniform estimates for finite truncations.

Let I be finite or countable, let $\mathcal{A}_i$ be arbitrary component sets, and let the cost and error functionals be finite and nonnegative. Assume the attainable diagrams

$$\Gamma_i = \{(s,t) \in [0,\infty)^2 : C_i(z_i) \le s,\ E_i(z_i) \le t \text{ for some } z_i \in \mathcal{A}_i\}$$

are closed, that each component contains an element of zero cost, and that

$$e_i = \min_{C_i(z_i)=0} E_i(z_i) < \infty, \qquad \inf_{z_i \in \mathcal{A}_i} E_i(z_i) = 0, \qquad M = \sum_i e_i < \infty.$$

Closedness gives a component minimizer for each finite constraint bound; approaching zero error may require unbounded cost. For $a, b \ge 2$, Theorem 7.4 constructs a countable model with $\sum_j W_j = M/b$, the maps and bounds of Theorem 1.1, and attained original and model minima. The maps are defined on admissible elements and binary vectors of finite cost. The admissible sets need not be discrete: the construction also applies to continuously parametrized families of approximants. Each model coordinate corresponds to an increment between selected points of a component diagram. Summability bounds the error from omitted increments, even when infinitely many components have positive cost and the total cost is finite.

For a finite coordinate set Q, define

$$\omega_Q = \sum_{j \notin Q} W_j, \qquad D_Q(B) = \omega_Q + \min_{\substack{S \subseteq Q \\ \sum_{j \in S} A_j \le B}} \sum_{j \in Q \setminus S} W_j.$$

Proposition 7.5 gives

$$0 \le D_Q(B) - D(B) \le \omega_Q, \qquad \max\{0, D_Q(B) - \omega_Q\} \le F(B) \le b D_Q(B/a). \tag{1.12}$$

A minimizer defining $D_Q(B/a)$ gives an admissible allocation of cost at most B and error at most $bD_Q(B/a)$. The term ω_Q is the error from the omitted coordinates; including it in D_Q gives the stated reconstruction estimate. As Q increases to the full coordinate set, $\omega_Q \to 0$, and the truncation estimate holds uniformly for $B \ge 0$. The comparison factors a, b remain unchanged as the finite models converge to D. Thus finite optimization gives admissible choices in the original countable problem with an explicitly bounded contribution from the omitted coordinates. Examples compute the weights, allocations, and error bounds for dyadic refinement, countably many approximation problems, and a continuous exponential error function.

For maximum error, replace summability by $\sup_i e_i < \infty$. Assume again that the component diagrams are closed, each component contains an element of zero cost, and each infimum error is zero. Theorem 8.4 gives, for every $a, b > 1$, positive weights A and a nonnegative model vector V such that

$$D_\infty(B) = E(B, V; \ell^0(A), \ell^\infty), \qquad D_\infty(B) \le F_\infty(B) \le b D_\infty(B/a), \tag{1.13}$$

where $F_\infty(B) = \inf_{\sum_i C_i(z_i) \le B} \sup_i E_i(z_i)$. Discarding selected coordinates after the first omitted coordinate preserves the maximum error and cannot increase cost. Proposition 8.5 gives uniform convergence of the finite model error functions when the supremum of the omitted model heights tends to zero. For normalized finite Pareto chains, Proposition 8.6 also gives $\widehat{D}_\infty(B) \le F_\infty(B) \le \widehat{D}_\infty(B/a)$ for every $a > 1$, with error factor one.

1.4 Exact additive and maximum-error allocation

Some allocation problems already possess the order that geometric selection creates in the general case. Their original increments give exact interpolation models. Consider component levels $0, \dots, m_i$, with nonnegative costs and errors and $C_i(0) = 0$, and put

$$a_{i,k} = C_i(k) - C_i(k-1), \qquad d_{i,k} = E_i(k-1) - E_i(k).$$

For finitely many rows, nonnegative nondecreasing marginal costs and nonnegative nonincreasing marginal gains give

$$\min_{\sum_i C_i(n_i) \le B} \sum_i E_i(n_i) = R_\Sigma + E(B, \mathbf{1}; \ell^0(a), \ell^1(d)), \qquad R_\Sigma = \sum_i E_i(m_i). \tag{1.14}$$

Theorem 3.1 replaces each selection by an initial segment without increasing either criterion. Equal marginal costs give nested optimal selections by sorting the gains, the classical allocation principle [5]. Theorem 3.4 extends the identity, with attained minima, to countable rows and levels when $\sum_i E_i(0) < \infty$ and costs diverge in each infinite row. The exchange replaces selected indices by earlier indices in the same row, whose costs are no larger and whose gains are no smaller. Telescoping identifies the cost and residual error of the resulting initial segment with those of an original component level.

For maximum error, assume only that $C_i(k)$ is nondecreasing and $E_i(k)$ is nonincreasing. With

$$h_{i,k} = E_i(k-1), \qquad R_\infty = \max_i E_i(m_i),$$

Theorem 8.1 gives

$$\min_{\sum_i C_i(n_i)\le B} \max_i E_i(n_i) = \max\{R_\infty, E(B,h;\ell^0(a),\ell^\infty)\}. \tag{1.15}$$

The first omitted coordinate determines the maximum error in its row. Setting all later coordinates to zero preserves this maximum, so the marginal costs and gains need not be ordered. The countable version has attained minima under $\sup_i E_i(0) < \infty$ and cost divergence in each infinite row.

1.5 Coordinate approximation in mixed sequence spaces

In coordinate approximation, the cost is the number of retained entries. In the matrix

$$f = \begin{pmatrix} 1 & 1 \\ 3/2 & 0 \end{pmatrix} \quad \text{in } \ell^2(\ell^1),$$

retaining a coefficient 1 gives squared error $13/4$, whereas retaining $3/2$ gives 4. The relevant gains are therefore the reductions in the total row errors.

For $0 < q \le p < \infty$, let $v_{i,j}$ be the decreasing row magnitudes and define

$$F_i(n) = \left(\sum_{j>n} v_{i,j}^q\right)^{p/q}, \qquad g_{i,n} = F_i(n-1) - F_i(n).$$

Let $\sigma_N(f)$ denote the smallest error in $\ell^p(\ell^q)$ obtained by retaining at most N entries of f. The gains decrease along each row. Their decreasing rearrangement (g_k^*) gives the exact error

$$\sigma_N(f)^p = \sum_{k>N} g_k^*. \tag{1.16}$$

When $p = q$, the gains are $v_{i,n}^p$, and the rule reduces to global magnitude thresholding. For $q < p$, the change in the whole row error governs the ordering. For finite or countable $f \in \ell^p(\ell^q)$, this gives nested optimal supports for every N. Such nesting can fail for every $p < q$. For all $0 < p, q < \infty$, Corollary 9.5 provides a weighted model with explicit error bounds for finite truncations. The convex hull construction below gives a nested family of approximate solutions for finite matrices.

1.6 Supporting lines and K–J representations

Linear penalties describe the supporting lines of the attainable diagram. For nonconvex data, a choice can minimize the error at a given cost bound without minimizing any positive linear penalty. The lower convex hull nevertheless determines every penalized minimum and leads to a single nested family of admissible approximations.

For normalized finite data, consecutive vertices of each component lower convex hull determine positive increments (c_β, d_β) with

$$\mathcal{K}(\lambda) = \sum_\beta \min(c_\beta, \lambda d_\beta),$$

which is an exact representation as a weighted ℓ^1 K-functional. Decompositions that assign each entire increment to one term, at the corresponding ratio of cost to error reduction, give an exact K–J representation. Grouping the increments on a geometric grid gives a single representation with an explicit bound in terms of the penalized minimum. Corollary 10.4 uses Cwikel–Jawerth–Milman's logarithmic averaging [3, Theorem 3.1] to obtain one integral representation in a weighted Banach pair, with separate endpoint bounds for every penalty parameter.

Ordering the ratios also gives a family of original admissible choices whose optimal error function H satisfies, for every $\alpha > 1$,

$$F(B) \le H(B) \le \frac{\alpha}{\alpha - 1} F(B/\alpha). \tag{1.17}$$

Each successive increment changes one component to another admissible choice. The ordering constructs the entire family without enumerating the product of the component choice sets. These factors occur in the weighted-sum approximation principle [8, Proposition 3 and Section 5]. Proposition 11.4 shows that uniform factors (α, β) for choices that minimize positive linear penalties require $\alpha^{-1} + \beta^{-1} \le 1$. The constructed family attains equality. Proposition 10.3 gives the sharp constant for disjoint binary decompositions on a geometric grid that assign each entire hull block to a single term.

Power penalties recover the information about constrained minima that linear penalties can omit. Set

$$\mathcal{K}_p(\lambda) = \inf_{z \in \mathcal{A}} \{C(z) + \lambda E(z)^p\}, \qquad p > 0.$$

Theorem 12.1 proves that, at every point of right-continuity of F,

$$F(B) = \lim_{p \to \infty} \left(\sup_{\lambda > 0} \frac{\mathcal{K}_p(\lambda) - B}{\lambda} \right)^{1/p}. \tag{1.18}$$

This includes every $B \ge 0$ for finite data and for the countable problems described above. The formula recovers the original error function, including portions lying above its lower convex envelope. For a sum of component errors the power is taken after summation; its computation therefore generally requires a nonseparable minimization.

1.7 Variational applications

A hierarchical decomposition provides corrections at several scales. One may choose which corrections to retain and at which accuracies to approximate them. Section 11.3 applies the allocation theorem to finite sets of such choices, using additive upper bounds for the norm of the approximant and its residual error. For binary retention, the resulting allocation is a knapsack problem. The construction applies, in particular, to corrections supplied by hierarchical methods for critical differential equations [16, 2].

Appendix A returns to three variational problems considered by Micchelli and Pinkus [10]: diagonal approximation, smoothing with a common constraint, and linear recovery from noisy data. The representation gives weighted models and maps back to admissible approximations, with bounds for both criteria. The diagonal and smoothing examples compute the weights explicitly. Their finite truncations give approximants and estimates for the error contributed by omitted coordinates.

1.8 Related work and the unpublished Paseky lectures

Pisier's formula for mixed ℓ^1/ℓ^∞ Banach couples [15] expresses the K-functional through component K-functionals whose parameters satisfy a common constraint. Here we minimize the total error under a bound on the sum of the component costs. The representation maps reconstruct elements of the prescribed admissible sets from binary vectors in the weighted model.

This paper publishes part of the material prepared for the authors' Paseky lectures in 2007 [9]. The Charles University booklet contained only an introduction; the full lecture manuscript remained unpublished. Chapter 5 of those notes interpreted the hierarchical decompositions of Tadmor, Nezzar, and Vese [17] through the K- and J-methods and the fundamental lemma. An independent solution of the L^2-convergence problem for hierarchical decompositions was given by Modin, Nachman, and Rondi [12].

Geometric approximation of attainable diagrams also has an extensive optimization literature. Papadimitriou and Yannakakis [13] establish small approximate Pareto sets and study their construction. Such sets consist of admissible elements representing the attainable cost–error pairs to a prescribed accuracy. The weighted models here provide an interpolation approximation functional and maps transferring individual choices in both directions. Section 11 constructs an explicit family of admissible elements by ordering the lower convex hull increments.

1.9 Organization of the paper

Sections 2–4 introduce the models, exact exchange principle, and geometric selection. Section 5 determines the sharp cost factor for replacement by initial segments, and Section 6 proves the finite variational fundamental lemma. Section 7 gives the countable construction, the bounds for all geometric parameters, and uniform estimates for finite truncations. Sections 8 and 9 treat maximum error and mixed sequence approximation.

Section 10 uses lower convex hulls to give exact representations of penalized minima and K–J representations. Section 11 then constructs a family of admissible approximate solutions and applies the allocation theorem to hierarchical corrections. Section 12 recovers the optimal error function from power penalties, and Section 13 discusses further questions. Appendix A develops the further variational applications.

2 Variational diagrams and interpolation models

2.1 The lower boundary and its supporting lines

For a finite admissible set $\mathcal{A}$ and nonnegative functionals C, E, define

$$\Gamma(C,E) = \{(t,u) \in [0,\infty)^2 : C(z) \le t,\ E(z) \le u \text{ for some } z \in \mathcal{A}\}. \tag{2.1}$$

This set is upward closed in both coordinates. Its lower boundary is described by the optimal error function

$$F(t) = \min_{C(z)\le t} E(z),$$

together with the vertical segments at its jumps. The minimum cost for a prescribed error bound is

$$C_*(u) = \min_{E(z)\le u} C(z).$$

An empty minimum has value $+\infty$. If each functional vanishes at some admissible element, both functions are finite on $[0,\infty)$.

The penalized functional is

$$\mathcal{K}(\lambda) = \min_z \{C(z) + \lambda E(z)\}, \qquad \lambda > 0.$$

The line $t + \lambda u = \mathcal{K}(\lambda)$ supports the diagram from below. The definitions give

$$\mathcal{K}(\lambda) = \inf_{B \geq 0} \{B + \lambda F(B)\}. \tag{2.2}$$

Indeed, a minimizer for $F(B)$ gives $\mathcal{K}(\lambda) \leq B + \lambda F(B)$. Conversely, for any z, the choice $B = C(z)$ gives $B + \lambda F(B) \leq C(z) + \lambda E(z)$.

One always has

$$\sup_{\lambda > 0} \frac{\mathcal{K}(\lambda) - B}{\lambda} \leq F(B). \tag{2.3}$$

The reverse inequality depends on convexity. Micchelli and Pinkus [10, Theorem 2.2 and Proposition 2.3] prove the corresponding relations for nonnegative convex functionals on a convex admissible set. For finite data, the left side of (2.3) is the lower convex envelope of the optimal error function. Section 10.4 gives examples of unsupported admissible elements that cannot be recovered from the penalized values.

2.2 Weighted binary vectors as an interpolation pair

Let J be finite and let $A_j, W_j > 0$. On the Abelian group $G = \{0,1\}^J$ with addition modulo two, define

$$|\xi|_A = \sum_j A_j \xi_j, \qquad |\xi|_W = \sum_j W_j \xi_j.$$

These are group norms: they are positive away from zero and invariant under negation. Subadditivity follows because the support of a sum is contained in the union of the two supports. Thus (G_A, G_W) is a pair of normed Abelian groups, for which interpolation was developed by Peetre and Sparr [14].

For $\mathbf{1} = (1)_{j \in J}$, subtraction in G takes the coordinate complement. The approximation and K-functionals of this element are therefore

$$E(B, \mathbf{1}; G_A, G_W) = \min_{\sum_j A_j \xi_j \leq B} \sum_j W_j (1 - \xi_j), \tag{2.4}$$

$$K(\lambda, \mathbf{1}; G_A, G_W) = \sum_j \min(A_j, \lambda W_j). \tag{2.5}$$

The second formula follows by minimizing independently in each binary coordinate.

There is an equivalent realization on real or complex vectors. Put

$$\|z\|_{\ell^0(A)} = \sum_{j : z_j \neq 0} A_j, \qquad \|w\|_{\ell^1(W)} = \sum_j W_j |w_j|,$$

and use the general approximation notation

$$E(B, f; U, V) = \inf_{\|z\|_U \leq B} \|f - z\|_V. \tag{2.6}$$

For $f = \mathbf{1}$, replacing each nonzero coordinate of z by one preserves the weighted support size and cannot increase the error. Hence (2.4) equals $E(B, \mathbf{1}; \ell^0(A), \ell^1(W))$. We use this model throughout the paper.

2.3 Comparison by separate bounds

The representation maps bound cost and error separately. These bounds compare constrained minima and, more generally, objectives that are nondecreasing in both variables.

Lemma 2.1 (Comparison of constrained and penalized minima)**.** *Let $\mathcal{A}$ and $\mathcal{X}$ be nonempty sets. Let C, E be functionals on $\mathcal{A}$, and C_d, E_d on $\mathcal{X}$, with C, C_d finite and nonnegative and E, E_d taking values in $[0,\infty]$. Suppose that maps $P\colon \mathcal{A}\to\mathcal{X}$ and $L\colon \mathcal{X}\to\mathcal{A}$ satisfy*

$$C_d(Pz)\le C(z), \qquad E_d(Pz)\le E(z), \qquad C(L\xi)\le aC_d(\xi), \qquad E(L\xi)\le bE_d(\xi)$$

for fixed $1\le a,b<\infty$. Define the optimal error functions by

$$F(B)=\inf_{C(z)\le B} E(z), \qquad D(B)=\inf_{C_d(\xi)\le B} E_d(\xi),$$

with an empty infimum equal to $+\infty$. Then, for every $B\ge 0$,

$$D(B)\le F(B)\le bD(B/a). \tag{2.7}$$

More generally, for every function $\Psi\colon [0,\infty)\times[0,\infty]\to[0,\infty]$ that is nondecreasing in each coordinate,

$$\inf_{\xi\in\mathcal{X}} \Psi(C_d(\xi),E_d(\xi)) \le \inf_{z\in\mathcal{A}} \Psi(C(z),E(z)) \le \inf_{\xi\in\mathcal{X}} \Psi(aC_d(\xi),bE_d(\xi)). \tag{2.8}$$

In particular, if

$$\mathcal{K}(\lambda)=\inf_z\{C(z)+\lambda E(z)\}, \qquad \mathcal{K}_d(\lambda)=\inf_\xi\{C_d(\xi)+\lambda E_d(\xi)\},$$

then

$$\mathcal{K}_d(\lambda)\le\mathcal{K}(\lambda)\le a\mathcal{K}_d((b/a)\lambda), \qquad \lambda>0. \tag{2.9}$$

The conclusions hold whether or not the infima are attained.

Proof. The map P sends admissible elements for $F(B)$ to admissible elements for $D(B)$. The map L sends admissible elements for $D(B/a)$ to admissible elements for $F(B)$ and increases the error by at most a factor b. Taking infima proves (2.7), also when one of the admissible sets is empty. Coordinatewise monotonicity gives

$$\Psi(C_d(Pz),E_d(Pz))\le\Psi(C(z),E(z)), \qquad \Psi(C(L\xi),E(L\xi))\le\Psi(aC_d(\xi),bE_d(\xi)).$$

Taking infima over z and ξ, respectively, proves (2.8). Set $\Psi(c,e)=c+\lambda e$ and factor out a to obtain (2.9). □

Applying these maps to the elements in the definitions of the diagrams gives

$$\operatorname{diag}(a,b)\Gamma_d\subseteq\Gamma\subseteq\Gamma_d.$$

Likewise, a model element with cost at most B and error at most η is reconstructed as an admissible element with cost at most aB and error at most $b\eta$. The fundamental lemma constructs these maps for each component. Adding the component estimates gives the same constants for independent component allocations. The comparison also applies to objectives formed from the total cost and error, such as $E(z)^{1/p}+\lambda C(z)^{1/q}$ for $p,q,\lambda>0$.

3 Exact knapsack allocation and the exchange principle

Suppose the admissible levels in row i are $0, 1, \dots, m_i$. Assume $C_i(0) = 0$, and define the marginal weights and gains by

$$a_{i,j} = C_i(j) - C_i(j-1), \qquad d_{i,j} = E_i(j-1) - E_i(j), \qquad 1 \le j \le m_i. \tag{3.1}$$

Write $r_i = E_i(m_i)$ for the terminal error in each row and $R_0 = \sum_i r_i$ for their sum. On the finite set $J = \{(i,j) : 1 \le i \le r,\ 1 \le j \le m_i\}$ put

$$A(z) = \sum_{(i,j):z_{i,j}\neq 0} a_{i,j}, \qquad D(w) = \sum_{(i,j)\in J} d_{i,j}|w_{i,j}|.$$

Theorem 3.1 (Exact allocation by weighted support size)**.** *Assume in each row that*

$$0 \le a_{i,1} \le \cdots \le a_{i,m_i}, \qquad d_{i,1} \ge \cdots \ge d_{i,m_i} \ge 0. \tag{3.2}$$

Then, for every $B \ge 0$,

$$\min_{\sum_i C_i(n_i)\le B} \sum_i E_i(n_i) = R_0 + E(B, \mathbf{1}; A, D). \tag{3.3}$$

The infimum in the approximation functional is attained. It is unchanged if restricted to binary vectors, or further to vectors of the form

$$z_{i,j} = \mathbf{1}_{\{j\le n_i\}}. \tag{3.4}$$

Zero marginal weights and zero marginal gains are allowed.

Proof. Given a real or complex competitor z, replace every nonzero coordinate by 1. This leaves its weighted support size unchanged and cannot increase $D(\mathbf{1} - z)$. It therefore suffices to consider a subset $S \subseteq J$, with total weight $\sum_{(i,j)\in S} a_{i,j}$ and residual error

$$R_0 + \sum_{(i,j)\notin S} d_{i,j}.$$

Suppose row i contains selected indices $j_1 < \cdots < j_{n_i}$. Since $j_k \ge k$, the monotonicity assumptions give

$$\sum_{k=1}^{n_i} a_{i,k} \le \sum_{k=1}^{n_i} a_{i,j_k}, \qquad \sum_{k=1}^{n_i} d_{i,k} \ge \sum_{k=1}^{n_i} d_{i,j_k}.$$

Replacing these indices by the first n_i indices of the row cannot increase the total weight or the residual error. After this replacement, telescoping gives

$$\sum_{j\le n_i} a_{i,j} = C_i(n_i), \qquad r_i + \sum_{j>n_i} d_{i,j} = E_i(n_i).$$

Thus every competitor can be replaced by initial segments without increasing either functional. These initial segments correspond precisely to allocations. There are finitely many binary vectors, so the minimum is attained. □

When the support functional is denoted by $\ell^0(a)$ and the error functional by $\ell^1(d)$, (3.3) reads

$$F(B) = R_0 + E(B, \mathbf{1}; \ell^0(a), \ell^1(d)).$$

When weights vanish, these functionals need not separate points. The formula also includes error reductions whose marginal weight is zero, since it involves no division by these weights.

Example 3.2 (Why the support functional matters)**.** Consider one item of weight 2 and gain 1, with constraint bound $B = 1$. The minimum residual error is 1, since the item cannot be selected. If weighted support size is replaced by the real weighted ℓ^1 expression $2|z|$, the vector $z = 1/2$ is admissible and gives residual $|1 - z| = 1/2$. The two functionals agree on binary vectors but give different unrestricted approximation problems.

3.1 Unit weights and nested optimal allocations

If $C_i(n) = n$, arrange all gains in decreasing order as $d_1^* \geq \cdots \geq d_M^* \geq 0$, $M = \sum_i m_i$. Break ties so that earlier entries in a row precede later entries in the same row.

Corollary 3.3 (Optimal allocations for unit weights)**.** *Under (3.2), with $C_i(n) = n$, for every integer $0 \leq N \leq M$,*

$$\min_{\sum_i n_i \leq N} \sum_i E_i(n_i) = R_0 + \sum_{k>N} d_k^*. \tag{3.5}$$

The first N gains determine an optimal allocation, and the corresponding sets of selected indices are nested as N increases.

Proof. Theorem 3.1 reduces the problem to maximizing the sum of at most N entries of a nonnegative list. The first N entries attain this maximum. Within each row the gains are nonincreasing, and the tie rule ensures that the selected entries form an initial segment. The selected sets are nested as N increases. □

A priority queue can store the next available gain in each row. Once the row gains have been computed, a binary heap makes N selections in $O(r + N \log(r + 1))$ operations. This is the classical marginal allocation procedure [5].

For unequal weights, the exact support representation still holds, but sorting by the ratio of gain to weight need not solve the constrained problem. Consider two rows, each with a single step, whose weights and gains are $(a, d) = (2, 3)$ and $(3, 4)$. With total weight bounded by 3, ordering by this ratio selects the first item, whereas the second gives the larger gain. The support representation retains the discrete constraint. The family of supported points in Section 11 provides another construction of allocations with explicit cost and error bounds.

3.2 Exact allocation with countably many rows and levels

When the original marginal weights and gains have the required monotonicity, the exact identity applies without geometric discretization. The following formulation extends the allocation theorem in the unpublished notes to infinitely many refinement levels [9, Theorem 9].

Theorem 3.4 (Exact allocation for countable data)**.** *Let I be finite or countable. In row i, let the admissible levels be $0, \ldots, m_i$, where m_i is a nonnegative integer or ∞; in an infinite row the allowed*

levels are precisely the nonnegative integers. Assume that $C_i(n), E_i(n)$ are finite and nonnegative, that $C_i(0) = 0$, and that the increments

$$a_{i,j} = C_i(j) - C_i(j-1), \qquad d_{i,j} = E_i(j-1) - E_i(j)$$

satisfy

$$a_{i,j} \geq 0, \qquad d_{i,j} \geq 0 \quad (1 \leq j \leq m_i),$$

and

$$a_{i,j} \leq a_{i,j+1}, \qquad d_{i,j} \geq d_{i,j+1} \quad (1 \leq j < m_i).$$

In an infinite row these conditions hold at every positive integer index. Suppose

$$\sum_{i \in I} E_i(0) < \infty, \qquad C_i(n) \longrightarrow \infty \quad \text{in every infinite row.} \tag{3.6}$$

Set $r_i = E_i(m_i)$ in a finite row and $r_i = \lim_{n\to\infty} E_i(n)$ in an infinite row, and put $R_0 = \sum_i r_i$. On $J = \{(i,j) : i \in I,\ 1 \leq j \leq m_i\}$ define

$$\|z\|_{\ell^0(a)} = \sum_{z_{i,j} \neq 0} a_{i,j}, \qquad \|w\|_{\ell^1(d)} = \sum_{(i,j)\in J} d_{i,j}|w_{i,j}|.$$

For every finite $B \geq 0$,

$$\min_{\sum_i C_i(n_i) \leq B} \sum_i E_i(n_i) = R_0 + E(B, \mathbf{1}; \ell^0(a), \ell^1(d)). \tag{3.7}$$

Both minima are attained. Every binary vector with finite weighted support size can be replaced by finite initial segments in each row without increasing either functional. The common value is right-continuous in B.

Proof. Nonnegative telescoping gives

$$\sum_{(i,j)\in J} d_{i,j} = \sum_i (E_i(0) - r_i) < \infty, \qquad E_i(n) = r_i + \sum_{j>n} d_{i,j}.$$

Replacing each nonzero coordinate of an approximant by 1 preserves its weighted support size and cannot increase its error. It therefore suffices to consider binary vectors. In an infinite row the nondecreasing marginal weights are eventually bounded below by a positive number: otherwise every increment would be zero, contrary to (3.6). Consequently a binary vector with finite weighted support size has only finitely many selected indices in that row. If these are $j_1 < \cdots < j_{n_i}$, then $j_k \geq k$, and

$$\sum_{k=1}^{n_i} a_{i,k} \leq \sum_{k=1}^{n_i} a_{i,j_k}, \qquad \sum_{k=1}^{n_i} d_{i,k} \geq \sum_{k=1}^{n_i} d_{i,j_k}.$$

Replacing the selected indices by the first n_i indices therefore does not increase the total weight or residual error. Summing these inequalities and using telescoping identifies selections of initial segments with allocations and proves equality of the two infima in (3.7).

For attainment, take a minimizing sequence of binary vectors with weighted support size at most B. Since J is countable, a diagonal subsequence converges at each coordinate to a binary vector ξ. Fatou's inequality gives

$$\sum_{i,j} a_{i,j}\xi_{i,j} \leq \liminf_{\nu} \sum_{i,j} a_{i,j}\xi_{i,j}^{(\nu)} \leq B.$$

Summability of d gives

$$\sum_{i,j} d_{i,j}(1-\xi_{i,j}^{(\nu)}) \longrightarrow \sum_{i,j} d_{i,j}(1-\xi_{i,j}).$$

Hence ξ attains the binary minimum. Replacing its row selections by initial segments gives an allocation attaining the same minimum. For right-continuity, take $B_\nu \downarrow B$ and binary minimizers for these bounds. A coordinatewise convergent subsequence has weighted support size at most B by Fatou's inequality and error equal to the limit of the minimum errors by summability of d. Monotonicity supplies the reverse inequality, so the limiting value is the minimum at B. □

The theorem allows selections in infinitely many rows with finite total weight. The divergence assumption on each infinite row ensures that a finite constraint bound permits only a finite refinement level in that row. If there are finitely many rows, finite initial errors already imply the summability assumption.

4 Geometric selection on the variational diagram

A general optimal error function may have large jumps and long constant intervals, and later refinements may give larger marginal gains. Gagliardo's coordinate selection chooses points on its completed graph at two independent geometric scales. When both geometric factors are at least two, the resulting increments have the monotonicity required by the exchange principle. Section 5 treats the remaining parameters and determines how much the weighted support size can increase when a selection is replaced by initial segments. The construction follows [6, 7] and [9, Sections 2.2 and 4.2]. We specify the finite endpoints for use in the representation theorem.

4.1 The completed lower boundary and its geometric selection

Consider a nontrivial finite Pareto chain

$$(c_0,e_0),\dots,(c_N,e_N), \qquad 0=c_0<c_1<\cdots<c_N=T, \quad e_0>e_1>\cdots>e_N=0. \tag{4.1}$$

Here Pareto means that no listed point is greater than or equal to another listed point in both coordinates. The optimal error function is

$$F_1(B)=\min\{e_j : c_j \le B\}, \qquad B\ge 0.$$

The completed lower boundary consists of the horizontal steps and the vertical segments joining their endpoints. In particular, (c_1,e_0) is the top of its first vertical segment.

Fix $a,b>1$. Set $t_0=0$ and

$$(t_1,u_1)=(c_1,e_0). \tag{4.2}$$

Given (t_k,u_k), proceed as follows.

1. If $at_k \ge T$, stop and set $m=k$. Append the terminal point $(t_{m+1},u_{m+1})=(T,0)$.
2. Otherwise put $x=at_k<T$. If $F_1(x)\le u_k/b$, choose

$$(t_{k+1},u_{k+1})=(x,F_1(x)).$$

3. If instead $F_1(x) > u_k/b$, choose

$$u_{k+1} = u_k/b, \qquad t_{k+1} = \min\{c_j : e_j \le u_{k+1}\}.$$

In the last case $t_{k+1} > x$, and the point lies on a vertical segment. Before the terminal point, successive first coordinates grow by a factor of at least a, and error coordinates decrease by a factor of at least b:

$$t_{k+1} \ge at_k, \qquad u_{k+1} \le u_k/b, \qquad 1 \le k < m. \tag{4.3}$$

At least one inequality is an equality, and every nonterminal error coordinate is positive. The construction terminates because $t_k \le T$, with

$$m \le 1 + \lfloor \log_a(T/c_1) \rfloor . \tag{4.4}$$

Using consecutive boundary points, define the model functionals by

$$C^d(k) = t_k, \qquad E^d(k) = u_{k+1}, \qquad 0 \le k \le m, \tag{4.5}$$

and the corresponding optimal error function by

$$D_1(B) = \min_{0\le k\le m,\ t_k\le B} u_{k+1}. \tag{4.6}$$

In particular, $D_1(0) = e_0$. We use the terminal error $u_{m+1} = 0$ but do not add the terminal cost coordinate $t_{m+1} = T$ as another model cost. This coordinate may coincide with t_m. The endpoint convention preserves the increment estimates proved below.

Theorem 4.1 (Finite Gagliardo discretization)**.** *The construction above satisfies, for every $B \ge 0$,*

$$\boxed{D_1(B) \le F_1(B) \le bD_1(B/a).} \tag{4.7}$$

Let

$$C_*(v) = \min\{c_j : e_j \le v\}, \qquad C_*^d(v) = \min\{t_k : u_{k+1} \le v\}, \qquad v \ge 0.$$

Then the minimum costs for prescribed error bounds satisfy

$$C_*^d(v) \le C_*(v) \le aC_*^d(v/b). \tag{4.8}$$

The displayed factors hold exactly for this finite construction; no $1+\varepsilon$ enlargement is needed.

Proof. Each constructed boundary point has the two properties

$$F_1(t_k) \le u_k, \qquad F_1(s) \ge u_k \quad (0 \le s < t_k). \tag{4.9}$$

For the first point, these properties follow from (4.2). For a point of the form $(x, F_1(x))$, they follow from monotonicity. When a point is chosen by an error threshold, its first coordinate is the first value at which that threshold is reached, which gives both inequalities.

If $t_k \le B < t_{k+1}$, with $0 \le k < m$, then (4.9) gives

$$D_1(B) = u_{k+1} \le F_1(B).$$

For $B \ge t_m$ this inequality follows from $D_1(B) = 0$.

For the upper bound, choose the largest $k \in \{0, \dots, m\}$ such that $t_k \le B/a$. If $k = 0$, then $D_1(B/a) = u_1 = e_0$, which bounds $F_1(B)$. If $k = m$, then $B \ge at_m \ge T$, so $F_1(B) = 0$. In the remaining case $1 \le k < m$, at least one of the two geometric inequalities is an equality. If $t_{k+1} = at_k$, then $B \ge t_{k+1}$ and

$$F_1(B) \le u_{k+1} = D_1(B/a).$$

If $u_{k+1} = u_k/b$, then $B \ge t_k$ and

$$F_1(B) \le u_k = bu_{k+1} = bD_1(B/a).$$

This proves (4.7) for every constraint bound, including the jump points.

For the inverse comparison, evaluate $D_1 \le F_1$ at $B = C_*(v)$. For its upper bound, evaluate $F_1(aB) \le bD_1(B)$ at $B = C_*^d(v/b)$. The finite minima are attained, so these substitutions prove (4.8). □

4.2 Monotonicity of the marginal weights

Set

$$A_n = t_n - t_{n-1}, \qquad W_n = u_n - u_{n+1}, \qquad 1 \le n \le m. \tag{4.10}$$

These weights are positive and satisfy

$$t_k = \sum_{n \le k} A_n, \qquad u_{k+1} = \sum_{n > k} W_n. \tag{4.11}$$

Lemma 4.2 (Bounds and monotonicity of the marginal weights)**.** *For all $1 \le n \le m$,*

$$(1 - a^{-1})t_n \le A_n \le t_n, \qquad (1 - b^{-1})u_n \le W_n \le u_n. \tag{4.12}$$

If $a, b \ge 2$, then $A_1 \le \cdots \le A_m$ and $W_1 \ge \cdots \ge W_m$. Hence both sequences in (4.5) are discretely convex as functions of the level index k.

Proof. The first estimate follows from $t_{n-1} \le t_n/a$ for $n \ge 2$; for $n = 1$ use $t_0 = 0$. The second follows from $u_{n+1} \le u_n/b$ for $n < m$; for $n = m$ use $u_{m+1} = 0$. If $a \ge 2$, then for $n < m$,

$$A_{n+1} \ge (a - 1)t_n \ge t_n \ge A_n.$$

If $b \ge 2$, then

$$W_n \ge (b - 1)u_{n+1} \ge u_{n+1} \ge W_{n+1}.$$

The increments of $E^d(k)$ are $-W_{k+1}$, which increase with k, so this error sequence is convex as asserted. □

The optimal error function $D_1(B)$ remains a step function. Theorem 3.1 uses convexity in the level index; the step function need not be continuous or convex in B.

Example 4.3 (An explicit discretization)**.** Take the admissible pairs

$$(0, 64), (1, 50), (3, 40), (7, 30), (9, 12), (16, 8), (25, 5), (40, 0).$$

For $a = b = 2$, the construction gives

$$(t_1, u_1) = (1, 64),\ (t_2, u_2) = (7, 32),\ (t_3, u_3) = (14, 12),\ (t_4, u_4) = (28, 5),$$

followed by $(40, 0)$. The model values and their increments are

Level k	$C^d(k)$	$E^d(k)$	A_k	W_k
0	0	64	—	—
1	1	32	1	32
2	7	12	6	20
3	14	5	7	7
4	28	0	14	5

Figure 1 shows the upper and lower bounds.

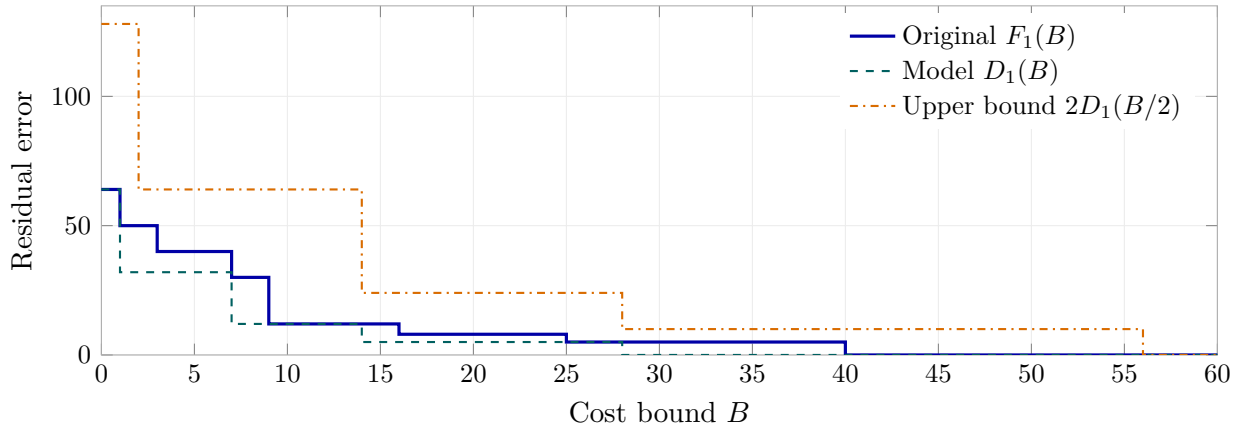


Figure 1: The model bounds the original optimal error function after rescaling the constraint bound and the error. Vertical segments complete the step graphs.

5 Knapsack models and sharp replacement by initial segments

Selecting arbitrary weighted items is a knapsack problem: an item of weight $A_{i,n}$ reduces error by $W_{i,n}$. A model allocation selects an initial segment in each row. When both geometric scales are at least two, the exchange principle shows that the two problems have the same value. We now compare them for every $a > 1$ and $b \geq 1$. Starting from the comparisons in [9, Section 4.4, Theorems 23 and 25], we determine the smallest factor by which the total weight may increase when arbitrary selections are replaced by initial segments without increasing the model error. Corollary 7.7 combines this replacement with the reconstruction map from the geometric model to bound the cost and error of the reconstructed admissible elements.

5.1 Extending a selection to initial segments

Let I be finite or countable. In each row let m_i be a positive integer or $+\infty$, and consider the geometric coordinates

$$t_{i,0} = 0 < t_{i,1} < t_{i,2} < \cdots, \qquad u_{i,1} \geq u_{i,2} \geq \cdots \geq 0.$$

In a finite row put $u_{i,m_i+1} = 0$; in an infinite row assume $u_{i,n} \to 0$. Fix $a > 1$ and $b \geq 1$, and assume

$$t_{i,n+1} \geq a t_{i,n} \quad (n < m_i), \qquad u_{i,n+1} \leq u_{i,n}/b \quad (n \leq m_i), \qquad M_d := \sum_i u_{i,1} < \infty. \tag{5.1}$$

For an infinite row the two inequalities hold at every positive integer index. For $b > 1$, a finite family of the rows constructed in Theorem 4.1 satisfies these assumptions. Lemma 7.3 gives the corresponding construction on a closed lower boundary for countable families with summable initial errors. In both constructions, the appended terminal cost coordinate is not an additional model cost.

Set

$$\Omega = \{(i,n) : i \in I,\ 1 \le n \le m_i\}, \qquad A_{i,n} = t_{i,n} - t_{i,n-1}, \qquad W_{i,n} = u_{i,n} - u_{i,n+1}.$$

The cost weights are positive, the error weights are nonnegative, and telescoping gives

$$\sum_{n \le k} A_{i,n} = t_{i,k}, \qquad \sum_{n > k} W_{i,n} = u_{i,k+1}, \qquad \sum_{(i,n)\in\Omega} W_{i,n} = M_d. \tag{5.2}$$

On

$$\Xi_A = \left\{ \xi \in \{0,1\}^\Omega : C_d(\xi) := \sum_{i,n} A_{i,n}\xi_{i,n} < \infty \right\}$$

write $E_d(\xi) = \sum_{i,n} W_{i,n}(1 - \xi_{i,n})$. Let Ξ_{pre} consist of those binary vectors whose selected indices in every row are $\{1, \dots, k_i\}$ for some finite $k_i \ge 0$. In a finite row $k_i \le m_i$. Define

$$D_{\mathrm{sel}}(B) = \inf_{\substack{\xi\in\Xi_A \\ C_d(\xi)\le B}} E_d(\xi) = E(B, \mathbf{1}; \ell^0(A), \ell^1(W)), \tag{5.3}$$

$$D_{\mathrm{pre}}(B) = \inf_{\substack{\xi\in\Xi_{\mathrm{pre}} \\ C_d(\xi)\le B}} E_d(\xi), \qquad B \ge 0. \tag{5.4}$$

For a finite family of constructed rows, D_{pre} is the optimal error function for the model allocation problem, denoted by D in Theorem 6.1. The two optimal error functions need not coincide when one of the geometric scales is below two. When some $W_{i,n}$ vanish, $\ell^1(W)$ in (5.3) denotes the associated error functional.

Proposition 5.1 (Completion to initial segments)**.** *Under (5.1), every $\xi \in \Xi_A$ can be extended to a selection $\widehat{\xi} \in \Xi_{\mathrm{pre}}$ of initial segments satisfying*

$$C_d(\widehat{\xi}) \le \frac{a}{a-1} C_d(\xi), \qquad E_d(\widehat{\xi}) \le E_d(\xi). \tag{5.5}$$

Consequently,

$$D_{\mathrm{sel}}(B) \le D_{\mathrm{pre}}(B) \le D_{\mathrm{sel}}\big((1 - a^{-1})B\big), \qquad B \ge 0. \tag{5.6}$$

The same completion bound holds without any geometric separation of the error coordinates, provided the weights $W_{i,n}$ are nonnegative and summable.

Proof. In an infinite row, $t_{i,n} \ge a^{n-1} t_{i,1}$ and $A_{i,n} \ge (1 - a^{-1}) t_{i,n} \to \infty$. A binary vector with finite weighted support size therefore selects only finitely many indices in each row. For a nonempty selected set S_i, put $k_i = \max S_i$ and select all earlier indices as well. Since

$$(1 - a^{-1}) t_{i,k_i} \le A_{i,k_i},$$

the total weight of the completed row satisfies

$$t_{i,k_i} \le \frac{a}{a-1} A_{i,k_i} \le \frac{a}{a-1} \sum_{n\in S_i} A_{i,n}.$$

If no index is selected in a row, leave its selection empty. Completion only adds selected indices, so it cannot increase the residual error. Summing over the finite or countable set of rows proves (5.5). Selections of initial segments are a subset of all selections, giving the first inequality in (5.6). Applying completion under the bound $(1 - a^{-1})B$ and taking infima gives the second. □

5.2 The optimal replacement factor

When $b \geq 2$, each error increment is at least the sum of all subsequent increments. This stronger form of monotonicity permits replacement by a shorter initial segment and improves the bound for total weight.

Theorem 5.2 (Sharp replacement by initial segments)**.** *For rows satisfying (5.1), define*

$$\rho(a,b) = \begin{cases} \dfrac{a}{a-1}, & 1 \leq b < 2, \\ \max\left\{1, \dfrac{1}{a-1}\right\}, & b \geq 2. \end{cases} \tag{5.7}$$

There is a map $\Pi\colon \Xi_A \to \Xi_{\mathrm{pre}}$ such that

$$C_d(\Pi\xi) \leq \rho(a,b) C_d(\xi), \qquad E_d(\Pi\xi) \leq E_d(\xi). \tag{5.8}$$

In particular,

$$D_{\mathrm{sel}}(B) \leq D_{\mathrm{pre}}(B) \leq D_{\mathrm{sel}}(B/\rho(a,b)), \qquad B \geq 0. \tag{5.9}$$

For each fixed $a > 1$ and $b \geq 1$, this factor is optimal over the stated class of rows, both for replacement by initial segments in (5.8) and for the comparison of optimal errors in (5.9). It is optimal even for a single finite row.

Proof. For $1 \leq b < 2$, use the completion map of Proposition 5.1. Suppose $b \geq 2$. In each row, leave the selection unchanged if it is already an initial segment. Otherwise, let j be the first missing index. At least one later index $k > j$ is selected. Replace the row selection by the initial segment ending at j.

All indices before j were already selected. Moreover,

$$W_{i,j} = u_{i,j} - u_{i,j+1} \geq (b-1)u_{i,j+1} \geq u_{i,j+1} = \sum_{n>j} W_{i,n}. \tag{5.10}$$

The gain from the newly selected index is therefore at least the total gain from the previously selected indices after j. Thus the replacement cannot increase the residual error. Its total weight is $t_{i,j}$, whereas the previously selected index k alone has weight

$$A_{i,k} = t_{i,k} - t_{i,k-1} \geq (a-1)t_{i,k-1} \geq (a-1)t_{i,j}.$$

Thus the new total weight in the row is at most $1/(a-1)$ times the previous total weight. An unchanged initial segment retains its weight. Taking the larger of the two factors gives the second case of (5.7). This defines Π separately in each row. Summing proves (5.8), also for countably many rows. Taking infima proves (5.9). Sharpness follows from Example 5.3 below. □

Example 5.3 (Extremal rows)**.** For $b > 1$, consider a row with two items and coordinates

$$(t_0, t_1, t_2) = (0, 1, a), \qquad (u_1, u_2, u_3) = (b, 1, 0). \tag{5.11}$$

Its weights are

$$(A_1, A_2) = (1, a-1), \qquad (W_1, W_2) = (b-1, 1),$$

and its three initial segments have cost–error pairs

$$(0, b), \qquad (1, 1), \qquad (a, 0).$$

Selecting only the second item gives the pair $(a-1, b-1)$. If $1 < b < 2$, an initial segment with error at most $b-1$ must include the whole row, with total weight a. The necessary ratio of weights is therefore $a/(a-1)$. If $b \geq 2$ and $1 < a < 2$, the initial segment containing only the first item has the least weight among those with error at most $b-1$; its weight is one, giving the necessary ratio $1/(a-1)$. For $a, b \geq 2$, a single-item row with $A_1 = W_1 = 1$ shows that the factor one is necessary.

These examples also prove that the bounds for optimal errors are sharp. In the first two cases, let $0 < r < \rho(a, b)$ and put $B = r(a-1)$. The binary vector selecting only the second item is admissible under the bound $B/r = a-1$, while every initial segment of weight at most B has error strictly larger than $b-1$. The single-item example gives the same conclusion for $r < 1$. At $b = 1$, replace $u_1 = b$ in (5.11) by any $\beta \in (1, 2)$. The error coordinates satisfy the required nonincreasing condition, and the argument for $1 < b < 2$ applies with β in place of b. Thus the factor $a/(a-1)$ is necessary also at $b = 1$.

For $b > 1$, the finite Gagliardo construction produces this two-item row. Choose $1 < \eta < b$ and $a < T \leq a^2$, and take the original Pareto chain

$$(0, b), \qquad (1, \eta), \qquad (a, 1), \qquad (T, 0).$$

The first point is $(t_1, u_1) = (1, b)$. With constraint bound a, the optimal error equals $1 = u_1/b$, so the second point is $(t_2, u_2) = (a, 1)$. The next step terminates because $at_2 = a^2 \geq T$. Hence the sharp factors occur within the geometric models constructed from finite admissible data.

The improved estimate also holds at $b = 2$: (5.10) still holds with equality possible. For example, when $a = 3/2$ and $b \geq 2$, completion gives factor 3, while optimal replacement by initial segments without increasing error gives factor 2.

Remark 5.4 (Comparison with the earlier integer bound). For $1 < a < 2$, define the integer used in [9, Section 4.4, Theorem 25] by

$$k(a) = \min\{k \in \mathbb{N} : a^k(1 - a^{-1}) \geq 1\}. \tag{5.12}$$

Then $k(a) \geq 2$ and

$$\frac{1}{a-1} \leq a^{k(a)-1}.$$

For $b \geq 2$, Theorem 5.2 therefore gives

$$D_{\mathrm{sel}}(B) \leq D_{\mathrm{pre}}(B) \leq D_{\mathrm{sel}}((a-1)B) \leq D_{\mathrm{sel}}(B/a^{k(a)-1}). \tag{5.13}$$

At $a = 3/2$, one has $k(a) = 3$ and $a^{k(a)-1} = 9/4$.

The threshold values $\alpha_k > 1$ are uniquely determined by

$$\alpha_k^{k-1}(\alpha_k - 1) = 1, \qquad k \geq 1. \tag{5.14}$$

They satisfy $\alpha_1 = 2$, $\alpha_k \downarrow 1$, and

$$\alpha_k - 1 \sim \frac{\log k}{k}, \qquad k(\alpha_k - 1) = \log k - \log\log k + o(1). \tag{5.15}$$

Indeed, the left side of (5.14) is strictly increasing in α_k, giving uniqueness and the decreasing sequence. Its limit must be one. Put $x_k = \alpha_k - 1$. The defining equation becomes

$$\log x_k + (k-1)\log(1+x_k) = 0.$$

For fixed $0 < \varepsilon < 1$, the left side evaluated at $x = (1 \pm \varepsilon)\log k/k$ equals $\pm\varepsilon \log k + \log\log k + O(1)$. Monotonicity therefore yields $x_k \sim \log k/k$. Substitution into the same equation, using $kx_k^2 \to 0$, gives $kx_k + \log(kx_k) = \log k + o(1)$ and the second assertion in (5.15).

6 The variational fundamental lemma and its consequences

Apply the geometric construction of Section 4 separately to r Pareto chains, using the same $a, b > 1$. Define

$$F(B) = \min_{\sum_i c_{i,j_i} \le B} \sum_i e_{i,j_i}, \tag{6.1}$$

$$D(B) = \min_{\sum_i t_{i,k_i} \le B} \sum_i u_{i,k_i+1}. \tag{6.2}$$

Adding the component estimates gives the same bounds for the total cost and error.

Theorem 6.1 (Representation of allocations with cost and error bounds)**.** *For every $B \ge 0$,*

$$\boxed{D(B) \le F(B) \le bD(B/a).} \tag{6.3}$$

Moreover, the reconstruction map sends each model allocation with total cost at most B to an original allocation with total cost at most aB and total error at most b times the model error.

If $a, b \ge 2$, let $A_{i,n}, W_{i,n}$ be the positive weights in (4.10), and let J_d be their finite index set. Then

$$D(B) = \min_{\substack{S \subseteq J_d \\ \sum_{j\in S} A_j \le B}} \sum_{j \notin S} W_j = E(B, \mathbf{1}; \ell^0(A), \ell^1(W)). \tag{6.4}$$

The selected items may always be chosen as initial segments within rows.

Proof. Map an original level j_i to a model minimizer for the constraint bound c_{i,j_i}; Theorem 4.1 bounds its cost and error by c_{i,j_i} and e_{i,j_i}. For the reconstruction, map a model level k_i to an original minimizer for the constraint bound at_{i,k_i}. Its cost is at most at_{i,k_i}, and

$$F_i(at_{i,k_i}) \le bD_i(t_{i,k_i}) = bu_{i,k_i+1}.$$

Summing these component bounds and applying Lemma 2.1 proves both (6.3) and the stated reconstruction estimate.

For $a, b \ge 2$, Lemma 4.2 gives the monotonicity of the marginal weights required by Theorem 3.1. The terminal errors vanish, so that theorem and (4.11) give (6.4). □

For $a = b = 2$, the representation gives

$$E(B, \mathbf{1}; \ell^0(A), \ell^1(W)) \le F(B) \le 2E(B/2, \mathbf{1}; \ell^0(A), \ell^1(W)). \tag{6.5}$$

Proof of Theorem 1.1. Discard dominated elements and choose one representative of each remaining cost–error pair. A component reduced to $(0,0)$ has no model coordinates: P_i is empty and L_i returns a fixed representative of $(0,0)$. The other components form Pareto chains as in (4.1); apply the geometric construction with the prescribed $a, b \geq 2$.

Use the component map defined in the preceding proof: choose a minimizer for $D_i(C_i(z_i))$ and let $P_i z_i$ be the indicator of its initial segment. Summing the bounds over the rows gives (1.4). To reconstruct from an arbitrary binary vector, replace its selected entries in each row by an initial segment of the same cardinality. By Lemma 4.2 and the exchange argument, neither model functional increases. Apply the component reconstruction map from the preceding proof to these initial segments and sum to obtain (1.5). Fixed rules for breaking ties in the finite admissible sets make the maps single-valued. The approximation identity and the bounds for optimal errors follow from Theorem 6.1. □

Corollary 6.2 (Transfer of approximate minimizers)**.** *Let a binary model vector satisfy $C_d(\xi) \leq B$ and $E_d(\xi) \leq D(B) + \varepsilon$. Then its reconstructed vector satisfies*

$$C(L\xi) \leq aB, \qquad E(L\xi) \leq bD(B) + b\varepsilon \leq bF(B) + b\varepsilon.$$

More generally, if $E_d(\xi) \leq \gamma D(B)$ for $\gamma \geq 1$, then $E(L\xi) \leq b\gamma F(B)$.

Proof. Apply (1.5) and the lower bound $D(B) \leq F(B)$. □

6.1 Normalization of general finite data

The elements of $\mathcal{A}_i$ need not be ordered initially. Set

$$c_i^{\min} = \min_{z \in \mathcal{A}_i} C_i(z), \qquad r_i^{\min} = \min_{z \in \mathcal{A}_i} E_i(z).$$

Subtract each constant from its respective coordinate and remove dominated pairs. Among pairs of equal cost, retain the one with the smallest error; among pairs of equal error, retain the one with the smallest cost. The remaining nontrivial pairs form a chain as in (4.1). A component reduced to $(0,0)$ contributes no increment weights.

Let $C_0 = \sum_i c_i^{\min}$, $R_0 = \sum_i r_i^{\min}$, and let $\widehat{F}, \widehat{D}$ be the optimal error functions for the normalized aggregate problems. Then for $B \geq C_0$,

$$F(B) = R_0 + \widehat{F}(B - C_0),$$

and the comparison becomes

$$R_0 + \widehat{D}(B - C_0) \leq F(B) \leq R_0 + b\widehat{D}((B - C_0)/a). \tag{6.6}$$

The minimum total error R_0 is added once, without multiplication by b.

6.2 Penalized values of the geometric model

Here $\mathcal{K}$ is computed from the normalized original chains, after subtracting the minimum costs and terminal errors. Define the corresponding model value by

$$\mathcal{K}_d(\lambda) = \min_{k_1, \ldots, k_r} \sum_i (t_{i,k_i} + \lambda u_{i,k_i+1}).$$

Lemma 2.1, applied to the maps of Theorem 6.1, gives

$$\mathcal{K}_d(\lambda) \le \mathcal{K}(\lambda) \le a\mathcal{K}_d((b/a)\lambda). \tag{6.7}$$

For $a, b \ge 2$, the penalized form of (6.4) separates over the items:

$$\mathcal{K}_d(\lambda) = \sum_{j \in J_d} \min(A_j, \lambda W_j). \tag{6.8}$$

Indeed, selecting an item contributes A_j to the objective, and leaving it unselected contributes λW_j. Because the weights are nondecreasing and the gains are nonincreasing along each row, the minimizing set of items can be chosen as an initial segment in each row. In particular, $a = b = 2$ yields

$$\mathcal{K}_d(\lambda) \le \mathcal{K}(\lambda) \le 2\mathcal{K}_d(\lambda).$$

6.3 Interpolation and approximation scales

The same model applies for all penalty parameters and constraint bounds. We can therefore compare the interpolation and approximation quantities obtained from the optimal value functions. For $0 < \theta < 1$ and $0 < q \le \infty$, set

$$\mathfrak{K}_{\theta,q}(\mathcal{K}) = \left\| \lambda^{-\theta} \mathcal{K}(\lambda) \right\|_{L^q((0,\infty), d\lambda/\lambda)}.$$

For $\sigma > 0$, similarly define

$$\mathfrak{E}_{\sigma,q}(F) = \left\| B^{\sigma} F(B) \right\|_{L^q((0,\infty), dB/B)}.$$

The usual supremum is used when $q = \infty$. These expressions are finite for normalized finite systems: F is bounded near zero and vanishes beyond a finite constraint bound, while $\mathcal{K}$ is bounded at infinity and is $O(\lambda)$ near zero.

Corollary 6.3 (Comparison of interpolation and approximation quantities)**.** *The model of Theorem 1.1 satisfies*

$$\mathfrak{K}_{\theta,q}(\mathcal{K}_d) \le \mathfrak{K}_{\theta,q}(\mathcal{K}) \le a^{1-\theta} b^{\theta} \mathfrak{K}_{\theta,q}(\mathcal{K}_d), \tag{6.9}$$

$$\mathfrak{E}_{\sigma,q}(D) \le \mathfrak{E}_{\sigma,q}(F) \le b a^{\sigma} \mathfrak{E}_{\sigma,q}(D). \tag{6.10}$$

Proof. The lower bounds follow from the corresponding pointwise inequalities. For the upper bound in (6.9), use (6.7) and substitute $s = (b/a)\lambda$. This gives the factor $a(b/a)^{\theta} = a^{1-\theta} b^{\theta}$. For (6.10), use $F(B) \le bD(B/a)$ and substitute $t = B/a$. The same substitutions apply to suprema. □

For the binary model, $\mathfrak{K}_{\theta,q}(\mathcal{K}_d)$ is the real interpolation quantity of $\mathbf{1}$ for the pair (G_A, G_W). The corresponding quantity on the original admissible set is defined by its variational diagram. Formula (6.9) compares the two, and the original admissible set need not have a vector space structure.

7 Infinite choices and countable allocations

This section treats countably many components, each of which may have infinitely many admissible elements. We extend the construction in the authors' unpublished Paseky notes [9, Sections 4.2–4.3, Lemmas 18 and 20, and Theorem 22] to closed component diagrams with summable initial errors. We prove attainment of the constrained minima and right-continuity of the optimal error function, then construct a weighted sequence model and maps between its binary vectors and the original admissible elements. Finite truncations of this model give explicit error estimates.

7.1 Closed component diagrams

Let I be a finite or countable index set. For each $i \in I$, let $\mathcal{A}_i$ be a nonempty set and let $C_i, E_i \colon \mathcal{A}_i \to [0,\infty)$ be finite-valued functions. Assume that the upward diagram

$$\Gamma_i = \{(t,u) \in [0,\infty)^2 : C_i(z_i) \le t,\ E_i(z_i) \le u \text{ for some } z_i \in \mathcal{A}_i\} \tag{7.1}$$

is closed, that C_i vanishes at some element of $\mathcal{A}_i$, and that

$$\inf_{z_i \in \mathcal{A}_i} E_i(z_i) = 0, \qquad e_i := \min_{C_i(z_i)=0} E_i(z_i), \qquad M := \sum_{i \in I} e_i < \infty. \tag{7.2}$$

The minimum defining e_i exists by closedness, as the next lemma shows. Fix a minimizer z_i^0 in the definition of e_i. No topology on $\mathcal{A}_i$ is assumed.

Lemma 7.1 (Component minima and right-continuity)**.** *For every $i \in I$ and $B \ge 0$, the minimum*

$$F_i(B) = \min_{C_i(z_i) \le B} E_i(z_i) \tag{7.3}$$

is attained. The function F_i is nonnegative, nonincreasing and right-continuous, with $F_i(0) = e_i$ and $F_i(B) \to 0$ as $B \to \infty$.

Proof. A zero-cost element is admissible for every $B \ge 0$. Let h be the infimum of E_i under the constraint $C_i \le B$; it is finite and nonnegative. For each $n \ge 1$, choose an admissible element with error less than $h + 1/n$. Then $(B, h+1/n) \in \Gamma_i$, and closedness gives $(B,h) \in \Gamma_i$. The definition of the diagram therefore provides an admissible element with error at most h. Its error must equal h.

Monotonicity follows from inclusion of the feasible sets. If $B_n \downarrow B$, the numbers $F_i(B_n)$ increase to a limit $h \le F_i(B)$. Since $(B_n, F_i(B_n)) \in \Gamma_i$, closedness gives $(B,h) \in \Gamma_i$, and therefore $F_i(B) \le h$. This proves right-continuity. Finally, C_i is finite-valued, so

$$\inf_{B \ge 0} F_i(B) = \inf_{z_i \in \mathcal{A}_i} E_i(z_i) = 0.$$

□

If $e_i = 0$, the fixed element z_i^0 satisfies $C_i(z_i^0) = E_i(z_i^0) = 0$. Such a component requires no model coordinates; the reconstruction map always returns z_i^0 in that component. We apply the geometric construction only to components with $e_i > 0$.

For the component data in (7.2), put

$$\mathcal{A}_C = \left\{ z \in \prod_{i \in I} \mathcal{A}_i : C(z) := \sum_i C_i(z_i) < \infty \right\}, \qquad E(z) = \sum_i E_i(z_i). \tag{7.4}$$

The error may be infinite at some elements of $\mathcal{A}_C$. The fixed zero-cost allocation $(z_i^0)_i$ has error $M < \infty$, so

$$F(B) = \inf_{\substack{z \in \mathcal{A}_C \\ C(z) \le B}} E(z)$$

is finite for every $B \ge 0$.

Lemma 7.2 (Attainment and right-continuity for the full allocation problem)**.** *Let I be finite or countable, let each $\mathcal{A}_i$ be nonempty, and let $C_i, E_i \colon \mathcal{A}_i \to [0,\infty)$ be finite-valued. Assume that each diagram (7.1) is closed, that each C_i vanishes somewhere, and that (7.2) holds, in particular $M = \sum_i e_i < \infty$. For every finite $B \geq 0$, the infimum defining $F(B)$ is attained by an element of $\mathcal{A}_C$. Moreover, F is nonnegative, nonincreasing and right-continuous on $[0,\infty)$, with $F(0) = M$ and $F(B) \leq M$ for every $B \geq 0$.*

Proof. The element $(z_i^0)_i$ is admissible for every $B \geq 0$ and has error M, so $0 \leq F(B) \leq M$. When $C(z) = 0$, nonnegativity forces $C_i(z_i) = 0$ in every component. Each error is then at least e_i, and $(z_i^0)_i$ attains their sum. Thus $F(0) = M$. The admissible sets increase with B, so F is nonincreasing.

We first show that each F_i is lower semicontinuous. If $b_\nu \to b \geq 0$ and $\varepsilon > 0$, then $b_\nu \leq b + \varepsilon$ for all sufficiently large ν. Monotonicity and Lemma 7.1 therefore give

$$\liminf_\nu F_i(b_\nu) \geq F_i(b+\varepsilon) \longrightarrow F_i(b) \quad \text{as } \varepsilon \downarrow 0.$$

Hence each F_i is lower semicontinuous, including at zero.

Fix $B \geq 0$ and choose admissible elements $z^{(\nu)}$ with $C(z^{(\nu)}) \leq B$ and $E(z^{(\nu)}) \to F(B)$. Each $C_i(z_i^{(\nu)})$ lies in $[0, B]$. Since I is finite or countable, a diagonal subsequence gives, after relabeling,

$$C_i(z_i^{(\nu)}) \longrightarrow \beta_i \quad \text{for every } i \in I.$$

Fatou's inequality gives

$$\sum_i \beta_i \leq \liminf_\nu \sum_i C_i(z_i^{(\nu)}) \leq B.$$

By lower semicontinuity and the definition of F_i,

$$F_i(\beta_i) \leq \liminf_\nu F_i(C_i(z_i^{(\nu)})) \leq \liminf_\nu E_i(z_i^{(\nu)}).$$

A second application of Fatou's inequality yields

$$\sum_i F_i(\beta_i) \leq \sum_i \liminf_\nu E_i(z_i^{(\nu)}) \leq \liminf_\nu E(z^{(\nu)}) = F(B).$$

For each i, choose a minimizer z_i for $F_i(\beta_i)$, which exists by Lemma 7.1. Their product z satisfies

$$C(z) \leq \sum_i \beta_i \leq B, \qquad E(z) = \sum_i F_i(\beta_i) \leq F(B).$$

Thus z belongs to $\mathcal{A}_C$ and satisfies the original constraint. The definition of the infimum gives $E(z) = F(B)$, proving attainment. The argument requires no topology on the sets $\mathcal{A}_i$.

To prove right-continuity, let $B_\nu \downarrow B$ and choose a minimizer $z^{(\nu)}$ for $F(B_\nu)$. Monotonicity implies that $F(B_\nu)$ increases to a finite limit $h \leq F(B)$. Each $C_i(z_i^{(\nu)})$ is bounded by B_1. The same diagonal argument gives limits β_i; Fatou's inequality and $B_\nu \to B$ give

$$\sum_i \beta_i \leq \liminf_\nu C(z^{(\nu)}) \leq B.$$

Lower semicontinuity of F_i and Fatou's inequality give $\sum_i F_i(\beta_i) \leq h$. Choosing minimizers for the $F_i(\beta_i)$ again gives an element z with $C(z) \leq B$ and error at most h. Hence

$$F(B) \leq h = \lim_\nu F(B_\nu) \leq F(B),$$

which proves right-continuity, including at zero. □

7.2 Geometric selection for infinitely many choices

For a finite Pareto chain, the construction starts at the smallest positive cost. An infinite admissible set need not have a least positive cost. We instead use the least cost bound that permits a prescribed error.

The starting point determines the sum of the error weights. For a finite Pareto chain, (4.2) starts with $u_1 = e$, and telescoping gives total weight e. Here we set $u_1 = e/b$ and use the least B for which error e/b is attainable. The initial model element then has error e/b and corresponds to the original element with $C = 0$ and error $e = bu_1$. The sum of the model error weights is therefore e/b. We use this convention for every component in this section, including those whose geometric construction terminates after finitely many points.

Lemma 7.3 (Geometric representation of a closed component diagram)**.** *Let F be the optimal error function of a component as in Lemma 7.1, with $F(0) = e > 0$, and fix $a, b > 1$. There is a finite or infinite sequence of positive boundary coordinates (t_k, u_k) such that, with $t_0 = 0$, the model choices*

$$C^d(k) = t_k, \qquad E^d(k) = u_{k+1}$$

satisfy

$$D_1(B) := \min_{t_k \le B} u_{k+1} \le F(B) \le bD_1(B/a), \qquad B \ge 0. \tag{7.5}$$

In the finite case the states have indices $0 \le k \le m$ and $u_{m+1} = 0$; in the infinite case they have indices $k \ge 0$.

The increments

$$A_n = t_n - t_{n-1}, \qquad W_n = u_n - u_{n+1} \tag{7.6}$$

are positive and satisfy

$$(1 - a^{-1})t_n \le A_n \le t_n, \qquad (1 - b^{-1})u_n \le W_n \le u_n, \tag{7.7}$$

$$t_k = \sum_{n \le k} A_n, \qquad u_{k+1} = \sum_{n > k} W_n, \qquad \sum_n W_n = e/b. \tag{7.8}$$

If $a, b \ge 2$, the sequence A_n is nondecreasing and W_n is nonincreasing. In an infinite row, $t_n \to \infty$ and $A_n \to \infty$.

Proof. Start with

$$u_1 = e/b, \qquad t_1 = \min\{B \ge 0 : F(B) \le e/b\}. \tag{7.9}$$

The set in this minimum is nonempty because $F(B) \to 0$. It is closed because F is nonincreasing and right-continuous. Right-continuity at zero and $F(0) > e/b$ give $t_1 > 0$. Put

$$T = \inf\{B \ge 0 : F(B) = 0\},$$

with $T = \infty$ when the set is empty. If $T < \infty$, then $F(T) = 0$ and $t_1 \le T$.

From a current point (t_k, u_k) proceed as follows. If $at_k \ge T$, stop, set $m = k$, and append the terminal point $(t_{m+1}, u_{m+1}) = (T, 0)$. Otherwise put $x = at_k < T$. When $F(x) \le u_k/b$, take

$$(t_{k+1}, u_{k+1}) = (x, F(x)).$$

When $F(x) > u_k/b$, take

$$u_{k+1} = u_k/b, \qquad t_{k+1} = \min\{B \ge 0 : F(B) \le u_{k+1}\}.$$

This threshold is finite and attained by the same argument as in (7.9); it is strictly larger than x. Every nonterminal error coordinate is positive. Between successive nonterminal points,

$$t_{k+1} \geq a t_k, \qquad u_{k+1} \leq u_k/b, \tag{7.10}$$

and at least one inequality is an equality.

If $T < \infty$, the construction terminates: the coordinates t_k are bounded by T and grow geometrically. As in the finite construction, the terminal coordinate T does not define an additional model element. If $T = \infty$, the construction continues for every k, and

$$t_k \geq a^{k-1} t_1 \longrightarrow \infty, \qquad 0 < u_k \leq e b^{-k} \longrightarrow 0. \tag{7.11}$$

In either case, only finitely many model indices satisfy $t_k \leq B$ for a given finite B, so the minimum defining $D_1(B)$ exists.

Each nonterminal point has the boundary properties

$$F(t_k) \leq u_k, \qquad F(s) \geq u_k \quad (0 \leq s < t_k). \tag{7.12}$$

At the initial point and at each threshold point, these properties follow from the minimality of the chosen first coordinate. At a point $(x, F(x))$, they follow from monotonicity. If $t_k \leq B < t_{k+1}$, the second property gives $D_1(B) = u_{k+1} \leq F(B)$. In the finite case, $D_1(B) = 0$ for the remaining values $B \geq t_m$, so the lower comparison holds everywhere.

For the upper comparison, let k be the largest model index with $t_k \leq B/a$. If $k = 0$, then

$$F(B) \leq e = b u_1 = b D_1(B/a).$$

If the row is finite and $k = m$, then $B \geq a t_m \geq T$ and both sides of the required upper bound are zero. In every other case $k \geq 1$ and the next nonterminal point exists. When $t_{k+1} = a t_k$, one has $B \geq t_{k+1}$ and

$$F(B) \leq F(t_{k+1}) \leq u_{k+1} = D_1(B/a).$$

When $u_{k+1} = u_k/b$, one has $B \geq t_k$ and

$$F(B) \leq F(t_k) \leq u_k = b u_{k+1} = b D_1(B/a).$$

This proves (7.5) at zero and at every discontinuity of F as well.

The increment estimates follow from $t_{n-1} \leq t_n/a$ and $u_{n+1} \leq u_n/b$, with $t_0 = 0$ and, in a finite row, $u_{m+1} = 0$. Telescoping and the zero limit of the error coordinates give (7.8). If $a \geq 2$, then

$$A_{n+1} \geq (a-1) t_n \geq t_n \geq A_n.$$

If $b \geq 2$, then

$$W_n \geq (b-1) u_{n+1} \geq u_{n+1} \geq W_{n+1}.$$

Finally, in an infinite row (7.11) and (7.7) imply $A_n \to \infty$. □

7.3 The countable representation

Theorem 7.4 (Representation for countably many components)**.** *Assume (7.1)–(7.2), with each Γ_i closed, and fix $a, b \geq 2$. There are a finite or countable set J, positive weights A_j, W_j with*

$$\sum_{j \in J} W_j = M/b, \tag{7.13}$$

and maps

$$P\colon \mathcal{A}_C \longrightarrow \Xi_A, \qquad L\colon \Xi_A \longrightarrow \mathcal{A}_C, \qquad \Xi_A = \left\{ \xi \in \{0,1\}^J : \sum_j A_j \xi_j < \infty \right\},$$

such that, on writing

$$C_d(\xi) = \sum_j A_j \xi_j, \qquad E_d(\xi) = \sum_j W_j (1 - \xi_j),$$

one has

$$C_d(Pz) \le C(z), \qquad E_d(Pz) \le E(z), \tag{7.14}$$

$$C(L\xi) \le a C_d(\xi), \qquad E(L\xi) \le b E_d(\xi). \tag{7.15}$$

The constrained minima for the original and model systems are attained, and

$$D(B) := \min_{\substack{\xi \in \Xi_A \\ C_d(\xi) \le B}} E_d(\xi) = E(B, \mathbf{1}; \ell^0(A), \ell^1(W)) \tag{7.16}$$

satisfies

$$\boxed{D(B) \le F(B) \le b D(B/a), \qquad B \ge 0.} \tag{7.17}$$

Both F and D are right-continuous. The model and the maps are independent of B.

Proof. Apply Lemma 7.3 to every row with $e_i > 0$, using the same a, b. Let J be the disjoint union of their increment index sets. This union is countable. Summing the nonnegative terms in (7.8) gives (7.13). For a row with $e_i = 0$, let P_i be the constant map whose value is the unique binary vector on the empty index set.

For $z \in \mathcal{A}_C$, choose in row i a model choice attaining $D_i(C_i(z_i))$, and let $P_i z_i$ be the indicator of the corresponding initial segment. The comparison for one component gives

$$C_{d,i}(P_i z_i) \le C_i(z_i), \qquad E_{d,i}(P_i z_i) \le F_i(C_i(z_i)) \le E_i(z_i).$$

Summing shows that $Pz \in \Xi_A$ and proves (7.14).

To define the reconstruction map L, consider $\xi \in \Xi_A$. Since $A_{i,n} \to \infty$ in each infinite row, only finitely many coordinates in that row can be nonzero. Replace them by the initial segment of the same cardinality. Lemma 7.3 shows that the $A_{i,n}$ are nondecreasing and the $W_{i,n}$ are nonincreasing, so this replacement increases neither C_d nor E_d, as in Theorem 3.1. If the resulting initial segment ends at k_i, choose an original minimizer for $F_i(a t_{i,k_i})$. It satisfies

$$C_i(L_i \xi_i) \le a t_{i,k_i}, \qquad E_i(L_i \xi_i) \le b D_i(t_{i,k_i}) = b u_{i,k_i+1}.$$

For an empty initial segment this is the fixed choice z_i^0, since $e_i = b u_{i,1}$. Rows with $e_i = 0$ also use z_i^0 and contribute no model coordinates. Summation proves (7.15); in particular $L\xi \in \mathcal{A}_C$ and $E(L\xi) \le M$. When a component has more than one minimizer, fix one of them. This makes P, L well-defined maps.

Lemma 7.2 proves attainment and right-continuity for the original problem. For the model, take a minimizing sequence of binary vectors satisfying $C_d \le B$. A diagonal subsequence converges in every coordinate to a binary vector ξ. Nonnegativity gives

$$C_d(\xi) \le \liminf_{\nu} C_d(\xi^{(\nu)}) \le B.$$

Because $\sum_j W_j < \infty$, coordinatewise convergence also gives

$$E_d(\xi^{(\nu)}) \longrightarrow E_d(\xi).$$

Indeed, convergence holds on every finite set of coordinates and the remaining error is uniformly bounded by the sum of the omitted W-weights. Thus ξ attains $D(B)$.

The maps give (7.17) by Lemma 2.1: apply P to a minimizer for $F(B)$ and L to a minimizer for $D(B/a)$. To prove (7.16), define

$$\|v\|_{\ell^0(A)} = \sum_{j:v_j\neq 0} A_j, \qquad \|w\|_{\ell^1(W)} = \sum_j W_j|w_j|.$$

The vector $\mathbf{1}$ belongs to $\ell^1(W)$ by (7.13). Replacing every nonzero coordinate of a competitor v by 1 preserves its weighted support size and does not increase its error. Thus the approximation infimum equals the minimum over binary vectors, whose attainment we have just proved.

Finally, let $B_\nu \downarrow B$ and choose a minimizer $\xi^{(\nu)}$ for $D(B_\nu)$. A diagonal subsequence converges at every coordinate to a binary vector ξ. Fatou's inequality gives

$$C_d(\xi) \le \liminf_\nu C_d(\xi^{(\nu)}) \le B,$$

so $\xi \in \Xi_A$ is admissible for $D(B)$. Summability of the W_j gives $E_d(\xi^{(\nu)}) \to E_d(\xi)$ by the same argument: convergence on finite sets and a uniform bound on the omitted terms. Since D is nonincreasing, $D(B_\nu)$ increases to a limit $h \le D(B)$. Consequently

$$D(B) \le E_d(\xi) = h = \lim_\nu D(B_\nu) \le D(B).$$

This proves right-continuity of D, including at zero, and completes the theorem. □

For $a = b = 2$, Theorem 7.4 gives the same comparison of optimal errors $D(B) \le F(B) \le 2D(B/2)$ as the finite theorem. The model may have infinitely many coordinates, while its total error weight is finite.

7.4 Finite truncations and admissible approximations

A binary vector supported on a finite set is zero at every omitted coordinate. The finite problem must therefore include the error contributed by those coordinates.

Proposition 7.5 (Error bounds for finite truncations)**.** *Let $Q \subseteq J$ be finite, set*

$$\omega_Q = \sum_{j\notin Q} W_j, \tag{7.18}$$

and define

$$D_Q(B) = \omega_Q + \min_{\substack{S\subseteq Q \\ \sum_{j\in S} A_j \le B}} \sum_{j\in Q\setminus S} W_j. \tag{7.19}$$

Then

$$0 \le D_Q(B) - D(B) \le \omega_Q \qquad (B \ge 0), \tag{7.20}$$

with the exact uniform error

$$\sup_{B\ge 0}\big(D_Q(B) - D(B)\big) = \omega_Q. \tag{7.21}$$

The original optimal error function satisfies the bounds

$$\max\{0, D_Q(B) - \omega_Q\} \le F(B) \le bD_Q(B/a), \qquad B \ge 0. \tag{7.22}$$

A minimizer in (7.19) with bound B/a, extended by zero outside Q and mapped by L, gives an original element z with $C(z) \le B$ and error at most $bD_Q(B/a)$.

Enumerate the components and retain the first m rows, with their first N_i model coordinates in row i. In a finite nontrivial row require $0 \le N_i \le m_i$, where m_i is its number of model coordinates. For a row with $e_i = 0$, set $N_i = 0$ and $u_{i,1} = 0$. For this choice of Q,

$$\omega_Q = \frac{1}{b} \sum_{i>m} e_i + \sum_{i \le m} u_{i,N_i+1} \tag{7.23}$$

$$\le \frac{1}{b} \left(\sum_{i>m} e_i + \sum_{i \le m} e_i b^{-N_i} \right). \tag{7.24}$$

Here a finite row included in its entirety has no omitted error, whereas omitting an entire row leaves error $u_{i,1} = e_i/b$.

Proof. The quantity $D_Q(B)$ is the minimum of the full model error over binary vectors supported in Q. Thus $D_Q(B) \ge D(B)$. If ξ attains $D(B)$, set its coordinates outside Q to zero. Its weighted support size cannot increase, and its error increases by

$$\sum_{j \notin Q} W_j \xi_j \le \omega_Q.$$

The truncated vector is admissible for $D_Q(B)$, proving (7.20) uniformly in B.

Summability of the weights also gives $D(B) \to 0$ as $B \to \infty$: given $\varepsilon > 0$, choose a finite set S with $\sum_{j \notin S} W_j < \varepsilon$ and take the indicator of S. Its weighted support size is the finite sum $\sum_{j \in S} A_j$. For $B \ge \sum_{j \in Q} A_j$, all retained coordinates can be selected and $D_Q(B) = \omega_Q$. Thus $D_Q(B) - D(B) \to \omega_Q$, and (7.20) proves (7.21).

Combine this inequality with $D(B) \le F(B) \le bD(B/a)$ to obtain (7.22). A minimizer for $D_Q(B/a)$ has full model error $D_Q(B/a)$, including the error at omitted coordinates. Applying (7.15) gives the asserted admissible element.

For a truncation to initial segments in each row, telescoping gives

$$\sum_{n > N_i} W_{i,n} = u_{i,N_i+1}, \qquad \sum_n W_{i,n} = e_i/b.$$

Summing these identities proves (7.23). The geometric decay in (7.11) gives $u_{i,N_i+1} \le e_i b^{-N_i-1}$ before termination; after termination the tail is zero. This proves (7.24). □

As finite sets Q increase to J, summability of the W_j gives $\omega_Q \to 0$. Thus the model's optimal error function is approximated uniformly on $[0, \infty)$ by finite constrained problems on binary vectors. Formula (7.24) separates the error due to omitted components from the error due to the truncation within each component.

We also have a direct estimate when the original family of components is truncated.

Corollary 7.6 (Truncation of the component family)**.** *For an enumeration of I, let $F^{[m]}(B)$ be the original minimum under the additional requirement $z_i = z_i^0$ for $i > m$. Then*

$$0 \le F^{[m]}(B) - F(B) \le \sum_{i>m} e_i, \qquad B \ge 0. \tag{7.25}$$

In particular, $F^{[m]} \to F$ uniformly on $[0, \infty)$.

Proof. The truncated minimum exists by the finite-component case of Lemma 7.2; its objective includes the constant error $\sum_{i>m} e_i$. Restricting the choices gives $F^{[m]}(B) \ge F(B)$. Take a minimizer for $F(B)$ and replace every component beyond m by z_i^0. Its cost cannot increase, and its error increases by at most $\sum_{i>m} e_i$, since the discarded errors are nonnegative. This proves (7.25). □

Both truncations preserve the bound on C. Summability of the initial errors gives the estimates without further assumptions on the functions C_i.

7.5 Penalized values and interpolation scales

For the infinite system define

$$\mathcal{K}(\lambda) = \inf_{z \in \mathcal{A}_C} \{C(z) + \lambda E(z)\}, \qquad \mathcal{K}_d(\lambda) = \sum_{j \in J} \min(A_j, \lambda W_j), \qquad \lambda > 0.$$

Both are finite: the element $(z_i^0)_i$ gives $\mathcal{K}(\lambda) \le \lambda M$, and $\mathcal{K}_d(\lambda) \le \lambda M/b$. The binary vector equal to one exactly when $A_j \le \lambda W_j$ has finite weighted support size, because

$$\sum_{A_j \le \lambda W_j} A_j \le \lambda \sum_j W_j < \infty.$$

It minimizes the penalized model functional. Minimizing separately in each coordinate gives the displayed sum. The maps (7.14)–(7.15) consequently yield

$$\mathcal{K}_d(\lambda) \le \mathcal{K}(\lambda) \le a\mathcal{K}_d((b/a)\lambda). \tag{7.26}$$

For the interpolation quantities defined in Section 6, the same changes of variables give, for $0 < \theta < 1$, $\sigma > 0$ and $0 < q \le \infty$,

$$\mathfrak{K}_{\theta,q}(\mathcal{K}_d) \le \mathfrak{K}_{\theta,q}(\mathcal{K}) \le a^{1-\theta} b^{\theta} \mathfrak{K}_{\theta,q}(\mathcal{K}_d), \tag{7.27}$$

$$\mathfrak{E}_{\sigma,q}(D) \le \mathfrak{E}_{\sigma,q}(F) \le b a^{\sigma} \mathfrak{E}_{\sigma,q}(D). \tag{7.28}$$

These inequalities hold with values in the extended nonnegative real numbers. Finiteness imposes additional regularity on the infinite data. Indeed, (7.26) uses the substitution $s = (b/a)\lambda$, producing $a(b/a)^{\theta}$, while (7.17) uses $t = B/a$, producing ba^{σ}. The argument applies to the corresponding suprema when $q = \infty$.

7.6 Transfer to original admissible choices

Combining Theorem 5.2 with the geometric constructions gives reconstruction maps for all parameters $a, b > 1$, both for finite admissible sets and for closed component diagrams.

Corollary 7.7 (Arbitrary geometric parameters)**.** *Apply the finite geometric construction to normalized finite component chains, or apply Lemma 7.3 to the finite or countable family of component error functions whose diagrams are closed and satisfy (7.2). Fix any $a, b > 1$, and use its weights in (5.3). There are fixed forward and reconstruction maps between $\mathcal{A}_C$ and Ξ_A such that*

$$C_d(Pz) \le C(z), \qquad E_d(Pz) \le E(z), \tag{7.29}$$

$$C(L\xi) \le a\rho(a,b)C_d(\xi), \qquad E(L\xi) \le bE_d(\xi). \tag{7.30}$$

Consequently, the original optimal error function satisfies

$$D_{\mathrm{sel}}(B) \le F(B) \le bD_{\mathrm{sel}}(B/(a\rho(a,b))), \qquad B \ge 0. \tag{7.31}$$

In particular, for $1 < a < 2$ and $b \ge 2$ the displayed factors for C and E are $a/(a-1)$ and b. For $a, b \ge 2$ they are a and b, as in the variational fundamental lemma. For $1 < b < 2$ they are $a^2/(a-1)$ and b; choosing $a = 2$ gives a cost factor of four with any error factor $1 < b < 2$.

Proof. Rows containing an element on which both C_i and E_i vanish contribute no model coordinates. Fix such an element in each of these rows; the reconstruction map returns it. For every original element z_i, minimize the component model with bound $C_i(z_i)$ and take the corresponding initial segment. Theorem 4.1 and Lemma 7.3 bound the resulting values of $C_{d,i}$ and $E_{d,i}$ by $C_i(z_i)$ and $E_i(z_i)$. This defines P and proves (7.29).

Given $\xi \in \Xi_A$, first apply the fixed map Π from Theorem 5.2, which replaces the nonzero coordinates in each row by an initial segment. If its length in row i is k_i, choose a minimizer for $F_i(at_{i,k_i})$. At this element, C_i is at most at_{i,k_i} and the error is at most bu_{i,k_i+1}. Summing and using (5.8) gives (7.30). For countably many rows, these bounds ensure that both C and E are finite. The chosen component minimizers therefore define an element of $\mathcal{A}_C$. Finally apply Lemma 2.1 to these two maps to obtain (7.31). □

The sharp constant $\rho(a, b)$ describes replacement by initial segments without increasing model error. The factors $(a\rho(a, b), b)$ in (7.30) are obtained by composing that replacement with the geometric reconstruction map. For $a = 3/2$ and $b \ge 2$, the resulting allocation has cost at most three times the weighted support size of the binary vector and error at most b times its residual error. More generally, the same fixed maps transfer every coordinatewise nondecreasing objective through Lemma 2.1.

7.7 Worked examples and the role of closedness

Example 7.8 (An infinite discrete problem and its finite truncations)**.** Consider one component with levels $n = 0, 1, \ldots$ and

$$C(n) = 2^n - 1, \qquad E(n) = 2^{-n}. \tag{7.32}$$

At each refinement level, the error is halved and the increment of C doubles. There are infinitely many admissible levels, none with zero error. Since $C(n)$ tends to infinity, only finitely many levels satisfy any fixed bound on C, and the upward diagram is closed. For $B \ge 0$ put

$$m(B) = \lfloor \log_2(B + 1) \rfloor .$$

The largest admissible level is $m(B)$, and therefore

$$F(B) = 2^{-m(B)}. \tag{7.33}$$

For $a = b = 2$, the initial threshold in (7.9) is $(t_1, u_1) = (1, 1/2)$. The construction then gives

$$t_k = 2^k - 1, \qquad u_k = 2^{-k}, \qquad k \geq 1. \tag{7.34}$$

Indeed, if these formulas hold at k, then $2t_k = 2^{k+1}-2$ is still smaller than the next value of C, $2^{k+1}-1$. Hence $F(2t_k) = 2^{-k} > u_k/2$, and the error-threshold step selects $(t_{k+1}, u_{k+1}) = (2^{k+1} - 1, 2^{-k-1})$. Thus the model has infinitely many positive coordinates, with

$$A_n = 2^{n-1}, \qquad W_n = 2^{-n-1}, \qquad n \geq 1. \tag{7.35}$$

On an initial segment of length n, C_d equals $2^n - 1$ and the error is 2^{-n-1}. The exchange argument shows that these initial segments are optimal, so

$$D(B) = 2^{-m(B)-1} = \frac{1}{2}F(B). \tag{7.36}$$

The first levels and their model counterparts are

Level n	Common cost	Original error	Model error
0	0	1	1/2
1	1	1/2	1/4
2	3	1/4	1/8
3	7	1/8	1/16
4	15	1/16	1/32

The maps are explicit. Map the original level n to the initial segment of length n. This preserves the value of C and halves the error. Conversely, every binary vector with finite weighted support size has only finitely many nonzero coordinates. Let n be their number and return the original level n. Replacing the nonzero coordinates by the first n increases neither C_d nor the model error. This reconstruction map therefore satisfies

$$C(L\xi) \leq C_d(\xi), \qquad E(L\xi) \leq 2E_d(\xi). \tag{7.37}$$

Thus the cost factor is one in this example, and (7.36) compares the two minima at the same B.

Now retain only $Q_N = \{1, \ldots, N\}$, where Q_0 is empty. The omitted weight and the optimal error in the finite model are

$$\omega_{Q_N} = 2^{-N-1}, \tag{7.38}$$

$$D_{Q_N}(B) = 2^{-\min\{m(B),N\}-1}. \tag{7.39}$$

To verify the second formula, the best retained initial segment has length $\min\{m(B), N\}$; its error in the full model includes the omitted error ω_{Q_N}. In particular,

$$D_{Q_N}(B) = D(B) \quad \text{for } 0 \leq B < 2^{N+1} - 1, \qquad \sup_{B\geq 0}(D_{Q_N}(B) - D(B)) = 2^{-N-1}.$$

The supremum follows also directly from $D_{Q_N}(B) = 2^{-N-1}$ for $B \geq 2^N - 1$, while $D(B) \to 0$. These formulas include $N = 0$: a model with no coordinates has constant error $1/2$.

For example, with $N = 3$ the retained support weights are $(1, 2, 4)$, the retained error weights are $(1/4, 1/8, 1/16)$, and the omitted error is $1/16$. Explicitly,

$$D_{Q_3}(B) = \frac{1}{16} + \min_{\substack{\xi\in\{0,1\}^3 \\ \xi_1+2\xi_2+4\xi_3\leq B}} \left\{\frac{1-\xi_1}{4} + \frac{1-\xi_2}{8} + \frac{1-\xi_3}{16}\right\}.$$

For $B = 3$, the initial segment of length two gives $D_{Q_3}(3) = D(3) = 1/8$ and reconstructs level 2, whose original error is $F(3) = 1/4$. For $B = 7$, all three retained coordinates give $D_{Q_3}(7) = D(7) = 1/16$ and reconstruct the optimal level 3, with original error $1/8$. For $B = 15$, level 4 is admissible in the full problem and gives $F(15) = 1/16$, whereas the truncated model still returns level 3, with error $1/8$. At this value of B, the two model errors differ by $1/16 - 1/32 = 1/32$.

More generally, the truncated model reconstructs exactly the level $\min\{m(B), N\}$. Its original error is $2D_{Q_N}(B)$, so

$$\sup_{B\geq 0}\bigl(2D_{Q_N}(B) - F(B)\bigr) = 2^{-N}. \tag{7.40}$$

Thus this infinite problem has an explicit weighted representation and finite families of admissible choices whose errors approximate the optimum uniformly for $B \geq 0$.

Example 7.9 (Countably many components with infinitely many levels)**.** For each $i \geq 1$, let the admissible levels be $n_i \in \mathbb{N}_0$, with

$$C_i(n_i) = 2^{n_i} - 1, \qquad E_i(n_i) = 2^{-i-n_i}. \tag{7.41}$$

The ith component has initial error 2^{-i}. Its kth refinement increases C_i by 2^{k-1} and halves the remaining error. There are infinitely many components, each with infinitely many levels. The total initial error is $\sum_i 2^{-i} = 1$, and no finite level has zero error. Only finitely many levels in each row satisfy any fixed bound on C_i, so its upward diagram is closed. The hypotheses of Theorem 7.4 therefore hold.

The geometric construction can be computed explicitly. Take $a = b = 2$. In row i, its first point is

$$t_{i,1} = 1, \qquad u_{i,1} = 2^{-i-1}.$$

If $t_{i,k} = 2^k - 1$ and $u_{i,k} = 2^{-i-k}$, then the bound $2t_{i,k} = 2^{k+1} - 2$ still permits no level beyond k. Its error is $u_{i,k}$, so the construction uses the next error threshold $u_{i,k}/2$, first reached at cost $2^{k+1} - 1$. Induction gives

$$t_{i,k} = 2^k - 1, \qquad u_{i,k} = 2^{-i-k}, \qquad k \geq 1.$$

Thus the countable model has the explicit weights

$$A_{i,k} = 2^{k-1}, \qquad W_{i,k} = 2^{-i-k-1}, \qquad i, k \geq 1, \qquad \sum_{i,k} W_{i,k} = \frac{1}{2}. \tag{7.42}$$

These weights give a stronger comparison than the general theorem. The first n_i model coordinates in row i have total weight $2^{n_i} - 1$ and leave model error 2^{-i-n_i-1}. Every binary vector with finite weighted support size can be replaced by initial segments in each row without increasing C_d or the error, since the support weights increase and the error weights decrease along each row. Consequently, if

$$F(B) = \min_{\sum_i (2^{n_i}-1)\leq B} \sum_i 2^{-i-n_i}, \qquad D(B) = E(B, \mathbf{1}; \ell^0(A), \ell^1(W)),$$

then

$$F(B) = 2D(B), \qquad B \geq 0. \tag{7.43}$$

Every support weight is at least one. A finite bound on C_d therefore permits only finitely many nonzero coordinates and finitely many components with positive levels. The components left at level zero contribute their summable initial errors.

For the first few values of B, the minimizing levels are as follows:

Bound B	Optimal nonzero levels $(n_1, n_2, \dots)$	$F(B)$
0	none	1
1	(1)	3/4
2	(1, 1)	5/8
3	(1, 1, 1)	9/16
4	(2, 1)	1/2
5	(2, 1, 1)	7/16
6	(2, 1, 1, 1)	13/32

Here the omitted levels are zero. To verify the table, note that C_i is independent of i. If $i < j$ but $n_i < n_j$, swapping the two levels decreases the error by

$$(2^{-i} - 2^{-j})(2^{-n_i} - 2^{-n_j}) > 0$$

and leaves C unchanged. Hence a minimizer may be chosen with $n_1 \geq n_2 \geq \cdots$. For $0 \leq B \leq 6$, no level can exceed two, since $C_i(3) = 7$. If k components have level two and the following ℓ components have level one, then $C = 3k + \ell$ and the total error is

$$\frac{1}{4} + 2^{-k-2} + 2^{-k-\ell-1}.$$

For each integer B in the table, minimizing this expression over $k, \ell \geq 0$ with $3k + \ell \leq B$ gives the displayed value. In particular, as B increases from three to four, the minimizing levels change from $(1, 1, 1)$ to $(2, 1, 0)$.

A finite approximation with explicit error bounds is obtained by restricting the model to the rectangle

$$Q_{m,N} = \{(i, k) : 1 \leq i \leq m,\ 1 \leq k \leq N\}.$$

Its omitted model error is

$$\omega_{m,N} = \sum_{(i,k) \notin Q_{m,N}} W_{i,k} = \frac{1}{2}\left[2^{-m} + (1 - 2^{-m})2^{-N}\right]. \tag{7.44}$$

Include this omitted error in the finite model by defining

$$D_{m,N}(B) = \omega_{m,N} + \min_{\substack{\xi \in \{0,1\}^{Q_{m,N}} \\ \sum_{(i,k) \in Q_{m,N}} 2^{k-1} \xi_{i,k} \leq B}} \sum_{(i,k) \in Q_{m,N}} 2^{-i-k-1}(1 - \xi_{i,k}).$$

Setting the coordinates outside $Q_{m,N}$ to zero does not increase C_d and adds at most $\omega_{m,N}$ to the error. Conversely, every binary vector in the finite model is admissible in the full model after extension by zero. Hence

$$0 \leq D_{m,N}(B) - D(B) \leq \omega_{m,N}.$$

Using (7.43), we obtain bounds for the same B

$$\max\{0, 2D_{m,N}(B) - 2\omega_{m,N}\} \leq F(B) \leq 2D_{m,N}(B), \qquad B \geq 0. \tag{7.45}$$

The upper bound is attained by an admissible allocation with positive levels only in the first m components and levels at most N in each:

$$2D_{m,N}(B) = \min_{\substack{0 \leq n_i \leq N\ (1 \leq i \leq m) \\ \sum_{i=1}^{m} (2^{n_i} - 1) \leq B}} \left(\sum_{i=1}^{m} 2^{-i-n_i} + 2^{-m} \right).$$

Thus the finite model gives an admissible allocation and bounds for the minimum in the infinite problem. For example, $m = N = 10$ uses 100 model coordinates and gives

$$\omega_{10,10} = 2^{-10} - 2^{-21} < 10^{-3}.$$

The optimal error in this finite model approximates D uniformly over all $B \geq 0$ with error less than 10^{-3}. The optimal error of the corresponding restricted original problem approximates F uniformly with error less than $2 \cdot 10^{-3}$.

Example 7.10 (A continuous problem and a finite binary model)**.** Let $\mathcal{A} = [0, \infty)$, $C(s) = s$, and $E(s) = \exp(-s)$. The optimal error function is $F(B) = \exp(-B)$, and its upward diagram is closed. The error tends to zero only as $s \to \infty$. For $a = b = 2$, the construction starts at $(\log 2, 1/2)$ and gives

$$t_0 = 0, \qquad t_k = 2^{k-1} \log 2, \qquad u_k = 2^{-2^{k-1}}, \qquad k \geq 1.$$

Indeed, $F(2t_k) = u_k^2 \leq u_k/2$, so every subsequent step doubles the first coordinate. The weights of the binary model are

$$A_1 = \log 2, \qquad A_n = 2^{n-2} \log 2 \quad (n \geq 2), \qquad W_n = 2^{-2^{n-1}} - 2^{-2^n} \quad (n \geq 1).$$

Their total error weight is $\sum_n W_n = 1/2$. Since the A_n are nondecreasing and the W_n decrease, every binary vector with finite weighted support size can be replaced by an initial segment without increasing C_d or the error. Consequently,

$$D(B) = u_{k+1} \quad \text{when } t_k \leq B < t_{k+1}, \qquad k \geq 0.$$

For the initial segment ending at k, choose the original parameter $s = 2t_k$. Its error is u_{k+1} for $k \geq 1$; the empty initial segment is mapped to $s = 0$, whose error is $1 = 2D(0)$. The first states are

k	t_k	$F(t_k)$	$D(t_k)$	$E(2t_k)$
0	0	1	1/2	1
1	$\log 2$	1/2	1/4	1/4
2	$2 \log 2$	1/4	1/16	1/16
3	$4 \log 2$	1/16	1/256	1/256

For example, with the original bound $B = 2 \log 2$, the general comparison reads $1/16 = D(B) \leq F(B) = 1/4 \leq 2D(B/2) = 1/2$. The model minimizer for $B/2$ retains only the first coordinate. Its reconstruction is $s = 2 \log 2$, which attains the original minimum.

Restricting the model to its first $N \geq 1$ coordinates leaves omitted error $\omega_N = u_{N+1} = 2^{-2^N}$. If $D_N = D_{\{1,\dots,N\}}$ includes this error as in (7.19), then

$$D_N(B) = \max\{D(B), 2^{-2^N}\}, \qquad \sup_{B \geq 0} \big(D_N(B) - D(B)\big) = 2^{-2^N}.$$

Thus four coordinates approximate the model's optimal error function with uniform error $2^{-16} = 1/65536$. These finite values give bounds for the original problem through (7.22).

Example 7.11 (A missing boundary choice)**.** Consider one component with the following pairs (C, E):

$$(0, 1), \qquad (1, 1/n), \quad n \geq 1.$$

Its upward diagram is not closed, because $(1,0)$ is a missing limit point. The infimum of the errors under the bound $C \le 1$ is zero, but no admissible choice has zero error. Consequently a model with an attained minimum $D(1) \le F(1) = 0$ cannot have a map into these original choices satisfying $E(L\xi) \le bE_d(\xi)$: its minimizing binary vector would have to map to an element with zero error. Passing to the closure adds the limiting pair $(1,0)$, so the resulting model represents the completed problem. A reconstruction into the original set requires a positive error tolerance. The closedness assumption in Theorem 7.4 ensures that the stated maps take values in the original admissible sets and satisfy the multiplicative error bound.

8 Minimizing the maximum error

We next minimize the largest component error under a bound on the sum of the C_i. The unpublished Paseky notes treated this problem together with minimization of the sum of errors [9, Theorems 10 and 27]. The corresponding interpolation endpoint is ℓ^∞. We use the errors before successive refinement levels as the target coordinates. In each row, the largest omitted coordinate then equals the error before the first omitted refinement. This gives an exact representation when C_i is nondecreasing and E_i is nonincreasing. We also extend the geometric construction to countably many components with bounded initial errors.

8.1 An exact maximum-error representation

Let I be finite or countable. In row i, let the admissible levels be $0, 1, \dots, m_i$, where $m_i \in \mathbb{N}_0 \cup \{\infty\}$; when $m_i = \infty$, the allowed levels are exactly $\mathbb{N}_0$. Assume that $C_i(n), E_i(n)$ are finite and nonnegative, $C_i(0) = 0$, and

$$C_i(n-1) \le C_i(n), \qquad E_i(n-1) \ge E_i(n), \qquad M_\infty := \sup_i E_i(0) < \infty. \tag{8.1}$$

In an infinite row assume also that $C_i(n) \to \infty$. Define

$$a_{i,k} = C_i(k) - C_i(k-1), \qquad h_{i,k} = E_i(k-1), \qquad 1 \le k \le m_i,$$

on $J = \{(i,k) : i \in I,\ 1 \le k \le m_i\}$. Thus the first target coordinate is $E_i(0)$, and the coordinate following an initial segment of length n is $E_i(n)$. Put

$$r_i = \begin{cases} E_i(m_i), & m_i < \infty, \\ \lim\limits_{n\to\infty} E_i(n), & m_i = \infty, \end{cases} \qquad R_\infty = \sup_i r_i.$$

Empty maxima and suprema are understood to be zero. The support functional is $\|z\|_{\ell^0(a)} = \sum_{(i,k):z_{i,k}\neq 0} a_{i,k}$; zero weights are permitted, as in Theorem 3.1.

Theorem 8.1 (Exact representation of the maximum error)**.** *Under (8.1), with $C_i(n) \to \infty$ in every infinite row, the following identity holds for every finite $B \ge 0$:*

$$\begin{aligned} F_\infty(B) &:= \min_{\substack{n_i \in \{0,\dots,m_i\} \\ \sum_i C_i(n_i) \le B}} \sup_i E_i(n_i) \\ &= \max\left\{ R_\infty, E(B, h; \ell^0(a), \ell^\infty) \right\}. \end{aligned} \tag{8.2}$$

Both minima are attained, and the corresponding optimal error functions are right-continuous. An approximation minimizer can be chosen by retaining an initial segment in each row. The increments $a_{i,k}$ need only be nonnegative; no ordering of these increments or of the successive error reductions is required.

For a tolerance $\tau \geq 0$, set

$$\mathcal{B}_\infty(\tau) = \sum_i \min\{C_i(n) : E_i(n) \leq \tau\}, \tag{8.3}$$

where an empty minimum is $+\infty$. Then

$$\begin{aligned} F_\infty(B) \leq \tau \quad &\Longleftrightarrow \quad \mathcal{B}_\infty(\tau) \leq B, \\ F_\infty(B) &= \min\left\{\tau \geq R_\infty : \sum_{(i,k):h_{i,k}>\tau} a_{i,k} \leq B\right\}. \end{aligned} \tag{8.4}$$

For finite data, one decreasing ordering of the entries $h_{i,k}$, with row order preserved at ties, gives a nested family of exact minimizers for all B: retain the longest initial segment of this ordering whose sum of support weights is at most B.

Proof. In the approximation problem, replace an approximant by h on its support and by zero elsewhere. This increases neither its weighted support size nor its error. Conversely, every binary vector ξ gives the approximant $h\xi$, whose weighted support size is at most $\sum_{i,k} a_{i,k}\xi_{i,k}$. Thus the approximation value equals

$$d_\infty(B) = \inf_{\substack{\xi\in\{0,1\}^J \\ \sum_{i,k} a_{i,k}\xi_{i,k}\leq B}} \sup_{i,k} h_{i,k}(1-\xi_{i,k}).$$

To prove attainment, take a minimizing sequence and pass to a diagonal subsequence that converges in every coordinate to a binary vector ξ. Fatou's inequality gives the required bound on its weighted support size. At each coordinate, the limiting residual is at most the lower limit of the maximum residuals. Taking the supremum over coordinates gives

$$\sup_{i,k} h_{i,k}(1-\xi_{i,k}) \leq \liminf_\nu \sup_{i,k} h_{i,k}(1-\xi^{(\nu)}_{i,k}).$$

The limiting vector therefore attains $d_\infty(B)$.

Now fix a binary vector with finite weighted support size. Keep every finite row whose coordinates are all one. In each remaining row, let $n_i + 1$ be the first zero coordinate and set all subsequent coordinates to zero. An infinite row cannot be entirely nonzero, since the sum of its support weights is $\lim_n C_i(n) = \infty$. The resulting support in that row is the initial segment of length n_i, and its weighted size does not increase. The largest residual remains $h_{i,n_i+1} = E_i(n_i)$, since each discarded target entry is at most this value. Replacing the support by initial segments therefore preserves the maximum error.

The weighted size of an initial segment is $\sum_{k\leq n_i} a_{i,k} = C_i(n_i)$. Its model residual in row i is $E_i(n_i)$ unless a finite row has been entirely selected, when the residual is zero and the original error is r_i. Consequently, for every binary vector whose support consists of initial segments,

$$\sup_i E_i(n_i) = \max\left\{R_\infty, \sup_{i,k} h_{i,k}(1-\xi_{i,k})\right\}.$$

This proves both inequalities in (8.2). Replacing a model minimizer by initial segments gives a minimizer for the original problem.

For right-continuity, let $B_\nu \downarrow B$ and take minimizers for $d_\infty(B_\nu)$. A diagonal subsequence again has a binary limit whose weighted support size is at most B. Its residual is at most $\lim_\nu d_\infty(B_\nu)$ by the same lower semicontinuity argument. Since d_∞ is nonincreasing,

$$d_\infty(B) \le \lim_\nu d_\infty(B_\nu) \le d_\infty(B).$$

Right-continuity of the original optimal error function follows from (8.2).

To attain tolerance τ, every row must reach a level with error at most τ. When such levels exist, the first minimizes C_i, because C_i is nondecreasing. Choosing these levels proves the first equivalence in (8.4). In the binary model, error at most τ requires exactly that every coordinate with $h_{i,k} > \tau$ be selected; the other coordinates can be left unselected. The limiting error in every original row must also be at most τ. This proves the second formula, including the strict inequality in its index set.

For finite data, take the longest initial segment of the decreasing ordering of h whose weighted size is at most B. If the next entry has value v, every binary vector with residual less than v must be one at that entry and all preceding entries. The sum of their support weights exceeds B. The chosen initial segment is therefore optimal. Including the unavoidable error R_∞ preserves optimality. Row order at ties makes each selected row an initial segment, and the selections are nested as B increases. □

For finite data, a single ordering therefore determines the optimal maximum error for every B. The quantities to order are the errors $E_i(k-1)$ themselves, since each component must meet the prescribed tolerance.

Example 8.2 (Reducing the largest remaining error)**.** Take two components with increments $a_{i,k} = 1$ and error lists

$$E_1 = (10, 9, 0), \qquad E_2 = (8, 0).$$

The target rows are $(10, 9)$ and (8). Their decreasing ordering gives the following minimizing levels and maximum errors:

B	0	1	2	3
(n_1, n_2)	$(0,0)$	$(1,0)$	$(2,0)$	$(2,1)$
$F_\infty(B)$	10	9	8	0.

For $B = 1$, refining the first component lowers the maximum error to 9. Refining the second leaves maximum error 10. The second choice gives the smaller sum of errors: 10, compared with 17 for the first choice. The first component has increasing successive error reductions $1, 9$, which are permitted by the maximum-error theorem. The least value of C that permits tolerance 8 is 2; any smaller tolerance requires $C = 3$.

8.2 Closed component diagrams and geometric models

For general admissible sets, we assume that the minimum errors at zero cost are bounded across components. Let I be finite or countable, let each $\mathcal{A}_i$ be nonempty, and let $C_i, E_i \colon \mathcal{A}_i \to [0, \infty)$ be finite-valued. Assume that each upward attainable diagram (7.1) is closed and each component has a zero-cost choice. Put

$$e_i = \min_{C_i(z_i)=0} E_i(z_i), \qquad M_\infty = \sup_i e_i < \infty, \tag{8.5}$$

and fix a minimizer z_i^0 in each row. On $\mathcal{A}_C$ from (7.4), define

$$E_\infty(z) = \sup_i E_i(z_i), \qquad F_\infty(B) = \inf_{\substack{z\in\mathcal{A}_C \\ C(z)\le B}} E_\infty(z). \tag{8.6}$$

An arbitrary element of $\mathcal{A}_C$ may have infinite maximum error. The fixed zero-cost allocation $(z_i^0)_i$ has error M_∞.

Lemma 8.3 (Attainment for a bounded maximum error)**.** *Under these assumptions, $F_\infty(B)$ is attained for every finite $B \ge 0$. The function F_∞ is nonnegative, nonincreasing and right-continuous, with $F_\infty(0) = M_\infty$ and $F_\infty(B) \le M_\infty$.*

Proof. Closedness implies that the component minimum $F_i(t)$ is attained for every finite t: the limit of points (t, u_ν) with errors $u_\nu \downarrow F_i(t)$ belongs to the component diagram. Likewise, if $t_\nu \downarrow t$, the points $(t_\nu, F_i(t_\nu))$ converge to a point in the diagram, proving right-continuity of F_i. A nonincreasing right-continuous function is lower semicontinuous: if $t_\nu \to t$ with $t_\nu \ge 0$, then for every $\varepsilon > 0$ it eventually satisfies $F_i(t_\nu) \ge F_i(t + \varepsilon)$, and $\varepsilon \downarrow 0$ gives the assertion.

The zero-cost allocation shows that the infimum is finite. Choose a minimizing sequence for $F_\infty(B)$ and pass to a diagonal subsequence for which $C_i(z_i^{(\nu)}) \to \beta_i$ in every component. Each component cost lies in $[0, B]$, and Fatou's inequality gives $\sum_i \beta_i \le B$. Lower semicontinuity yields

$$F_i(\beta_i) \le \liminf_\nu F_i(C_i(z_i^{(\nu)})) \le \liminf_\nu E_i(z_i^{(\nu)}) \le \lim_\nu E_\infty(z^{(\nu)}) = F_\infty(B).$$

Choose a minimizer for each $F_i(\beta_i)$. Their product has cost at most B and maximum error at most $F_\infty(B)$, so it attains the infimum.

For right-continuity, take $B_\nu \downarrow B$ and minimizers for $F_\infty(B_\nu)$. Their maximum errors increase to a limit $h \le F_\infty(B)$. The same diagonal argument, with the component costs now bounded by B_1, gives $\sum_i \beta_i \le B$ and $F_i(\beta_i) \le h$ for every i. Minimizers for the $F_i(\beta_i)$ give $F_\infty(B) \le h$, proving equality. Finally, when the total cost is zero, every component cost is zero. Thus the least maximum error for $B = 0$ is $\sup_i e_i = M_\infty$. □

Theorem 8.4 (Geometric representation of maximum error)**.** *Assume that the component diagrams are closed, that each component has a zero-cost choice, and that (8.5) holds. Suppose also that $\inf_{z_i\in\mathcal{A}_i} E_i(z_i) = 0$ for every i. For every $a, b > 1$ there are a finite or countable index set J, positive weights A_j, V_j, and maps*

$$P\colon \mathcal{A}_C \longrightarrow \Xi_A, \qquad L\colon \Xi_A \longrightarrow \mathcal{A}_C, \qquad \Xi_A = \{\xi \in \{0,1\}^J : \sum_j A_j\xi_j < \infty\},$$

such that, on putting

$$C_d(\xi) = \sum_j A_j\xi_j, \qquad E_{\infty,d}(\xi) = \sup_j V_j(1-\xi_j), \tag{8.7}$$

one has

$$C_d(Pz) \le C(z), \qquad E_{\infty,d}(Pz) \le E_\infty(z), \tag{8.8}$$

$$C(L\xi) \le aC_d(\xi), \qquad E_\infty(L\xi) \le bE_{\infty,d}(\xi). \tag{8.9}$$

The weights can be chosen with $\sup_j V_j = M_\infty/b$. *The model minimum is attained and satisfies*

$$D_\infty(B) := \min_{C_d(\xi)\le B} E_{\infty,d}(\xi) = E(B, V; \ell^0(A), \ell^\infty),$$
$$\boxed{D_\infty(B) \le F_\infty(B) \le bD_\infty(B/a), \qquad B \ge 0.} \tag{8.10}$$

Both optimal error functions are right-continuous. The weights and maps are independent of B.

Proof. For each $e_i > 0$, apply the component construction of Lemma 7.3, with parameters $a, b > 1$. Closedness and $\inf E_i = 0$ give all the required properties of its optimal error function. Write the selected points as $(t_{i,n}, u_{i,n})$, and define

$$A_{i,n} = t_{i,n} - t_{i,n-1}, \qquad V_{i,n} = u_{i,n}. \tag{8.11}$$

After an initial segment of length k, the supremum of the remaining target entries is $u_{i,k+1}$. It is zero when a finite row is retained in full. These are the model errors in (7.5). The initialization $u_{i,1} = e_i/b$ gives $\sup_{i,n} V_{i,n} = M_\infty/b$. Rows with $e_i = 0$ contribute no coordinates; in these rows we use z_i^0.

For an original element z_i, minimize the component model under the bound $C_i(z_i)$ and take the corresponding initial segment. The component comparison gives model cost at most $C_i(z_i)$ and model error at most $E_i(z_i)$. Summing the costs and taking the supremum of the errors proves (8.8).

To define the reconstruction map, truncate each row of $\xi \in \Xi_A$ at its first zero coordinate. Since $t_{i,n} \to \infty$ in an infinite row, finite total cost excludes retaining that entire row. The argument in Theorem 8.1 shows that the retained initial segments have no larger cost and the same maximum error. If the initial segment in row i has length k_i, choose a minimizer for $F_i(at_{i,k_i})$. The component comparison gives

$$C_i(L_i\xi_i) \le at_{i,k_i}, \qquad E_i(L_i\xi_i) \le bu_{i,k_i+1}.$$

For the empty initial segment take z_i^0, whose cost is zero and whose error is $e_i = bu_{i,1}$. Summing the costs and taking the supremum of the errors proves (8.9). Fixing one minimizer for each of these component problems defines the maps.

The model levels in row i have costs $t_{i,k}$ and errors $u_{i,k+1}$. The costs increase and the errors decrease to zero; in every infinite row the costs tend to infinity. The initial errors are bounded by M_∞/b. Theorem 8.1 therefore gives attainment and right-continuity for the model, as well as the identity with the approximation functional in (8.10). Lemma 8.3 gives attainment and right-continuity for the original problem. The bounds for P and L then give the comparison of optimal errors by Lemma 2.1. □

The weights in this model are the heights $u_{i,n}$ themselves. The argument at the first missing coordinate applies for every $a, b > 1$. The ordering of the increments required for the sum of errors in Theorem 7.4 is unnecessary here.

8.3 Error bounds for finite truncations

Proposition 8.5 (Truncation by the largest omitted error)**.** *For the model of Theorem 8.4, let* $Q \subset J$ *be finite and put*

$$\eta_Q = \sup_{j\notin Q} V_j, \qquad D_{\infty,Q}(B) = \min_{\substack{S\subseteq Q \\ \sum_{j\in S} A_j \le B}} \max\left\{\eta_Q, \max_{j\in Q\setminus S} V_j\right\}. \tag{8.12}$$

Then, for every $B \geq 0$,

$$D_{\infty,Q}(B) = \max\{D_\infty(B), \eta_Q\}, \qquad 0 \leq D_{\infty,Q}(B) - D_\infty(B) \leq \eta_Q. \tag{8.13}$$

Extending a minimizer for $D_{\infty,Q}(B/a)$ by zero and applying L gives an original element $z_{B,Q}$ of cost at most B, with

$$\begin{aligned}(D_{\infty,Q}(B) - \eta_Q)_+ &\leq F_\infty(B) \leq E_\infty(z_{B,Q}) \\ &\leq bD_{\infty,Q}(B/a) \leq b\max\{F_\infty(B/a), \eta_Q\}.\end{aligned} \tag{8.14}$$

Only finitely many components of this choice differ from the fixed zero-cost representatives.

For $I = \mathbb{N}$, retain the first N_i model coordinates in each of the first m rows, with N_i a nonnegative integer. In a finite nontrivial model row of length m_i, require $0 \leq N_i \leq m_i$. Take $N_i = 0$ and $u_{i,1} = 0$ for a row with $e_i = 0$. Then

$$\begin{aligned}\eta_Q &= \max\left\{b^{-1}\sup_{i>m} e_i, \max_{i\leq m} u_{i,N_i+1}\right\} \\ &\leq b^{-1}\max\left\{\sup_{i>m} e_i, \max_{i\leq m} e_i b^{-N_i}\right\}.\end{aligned} \tag{8.15}$$

An entirely retained finite row has remaining height zero. If $e_i \to 0$, increasing finite sets can therefore be chosen so that $\eta_Q \to 0$. The optimal error functions of the finite models then converge uniformly for $B \geq 0$.

Proof. Restricting binary vectors to Q gives $D_{\infty,Q}(B) \geq D_\infty(B)$, and the omitted coordinates force $D_{\infty,Q}(B) \geq \eta_Q$. Conversely, truncate a full minimizer for $D_\infty(B)$ to Q. Its cost cannot increase, and its maximum residual is at most $\max\{D_\infty(B), \eta_Q\}$. This proves the equality in (8.13) and its error bound. Extending a finite optimizer by zero and applying L gives its cost and error bounds in (8.14); the remaining inequalities follow from (8.10) and (8.13). A finite support meets only finitely many rows. In every other row, the chosen representative is z_i^0.

In an omitted row the largest height is $u_{i,1} = e_i/b$; in a partially retained row it is u_{i,N_i+1}. These observations give the equality in (8.15). The geometric bound $u_{i,n} \leq e_i b^{-n}$ gives its inequality. When $e_i \to 0$, first choose enough rows that the supremum over omitted rows is small, then retain enough coordinates in each of the remaining rows. The resulting $\eta_Q \to 0$ and (8.13) prove uniform convergence. □

Bounded initial errors suffice for the representation and attainment. If the omitted heights tend to zero, the finite optimal error functions converge uniformly. Bounded initial errors alone do not imply this convergence. For example, let every row have two states, with $C_i(1) = 2^{-i}$, $E_i(0) = 1$, and $E_i(1) = 0$. The geometric model has one coordinate per row, of cost 2^{-i} and height $1/b$. All coordinates can be retained under the bound $C_d \leq 1$, so $D_\infty(1) = 0$, whereas every finite restriction has omitted height $1/b$ and $D_{\infty,Q}(1) = 1/b$.

8.4 Discretizing cost without increasing error

For finite Pareto chains, discretizing only the cost coordinate gives an error factor of one. We thus obtain a finite model at the endpoint $b = 1$.

Proposition 8.6 (Discretization of the cost coordinate)**.** *Let the original choices be finitely many normalized finite Pareto chains as in (4.1), and fix $a > 1$. There is a finite maximum-error support model $\widehat{D}_\infty$ and maps with separate factors $(a, 1)$ such that*

$$\widehat{D}_\infty(B) \le F_\infty(B) \le \widehat{D}_\infty(B/a), \qquad B \ge 0. \tag{8.16}$$

The reconstruction of a minimizer for $\widehat{D}_\infty(B/a)$ therefore has cost at most B and maximum error at most $\widehat{D}_\infty(B/a)$.

Proof. Omit a row with a zero-cost, zero-error choice, retaining that fixed choice for the reconstruction map. In a nontrivial row let $c_1 > 0$ be its first positive cost on the Pareto chain, T its first zero-error cost, and $e = F_i(0)$. Set $t_0 = 0$ and $t_k = a^{k-1}c_1$ for $1 \le k \le m$, stopping at the first m for which $at_m \ge T$. There are finitely many selected costs, and

$$m \le 1 + \lfloor \log_a(T/c_1) \rfloor.$$

Give the model state k cost t_k and error

$$q_0 = e, \qquad q_k = F_i(at_k) \quad (1 \le k \le m).$$

The sequence q_k is nonincreasing and $q_m = 0$. Let $\widehat{D}_i(B) = \min_{t_k \le B} q_k$. For $B < c_1$, both $\widehat{D}_i(B)$ and $F_i(B)$ equal e. For $t_k \le B < t_{k+1} = at_k$, the model error is $F_i(at_k) \le F_i(B)$, and beyond its last selected cost the model error is zero. Hence $\widehat{D}_i(B) \le F_i(B)$ everywhere. If t_k is the largest model cost at most B/a, then for $k \ge 1$,

$$F_i(B) \le F_i(at_k) = q_k = \widehat{D}_i(B/a),$$

while $k = 0$ gives $F_i(B) \le e = q_0$. This proves the row comparison with factors $(a, 1)$.

Use support weights $\widehat{A}_{i,k} = t_k - t_{k-1}$ and target entries $\widehat{V}_{i,k} = q_{k-1}$. Theorem 8.1 identifies this component allocation problem with the finite support model. Map each original element to a minimizer of the component model under the same cost bound. Conversely, truncate a binary vector at the first zero coordinate in each row. Map an initial segment of length k to a minimizer for $F_i(at_k)$. It has cost at most at_k and error q_k; for $k = 0$, use the fixed zero-cost element. Summing the costs and taking the supremum of the errors proves the asserted bounds for the maps. Lemma 2.1 then gives (8.16). □

9 Best coordinate approximation in mixed sequence spaces

For a finite matrix $f = (f_{i,j})$ and $0 < p, q < \infty$, define

$$\|f\|_{\ell^p(\ell^q)} = \left[\sum_i \left(\sum_j |f_{i,j}|^q \right)^{p/q} \right]^{1/p}.$$

Its best coordinate approximation error is

$$\sigma_N(f) = \inf_{\#\operatorname{supp} h \le N} \|f - h\|_{\ell^p(\ell^q)}.$$

Within each row arrange the magnitudes as

$$v_{i,1} \ge \cdots \ge v_{i,m_i} \ge 0.$$

On a prescribed support, an optimal approximant reproduces the selected coefficients. If n_i coefficients are retained in a row, we may choose the largest n_i magnitudes. Thus

$$\sigma_N(f)^p = \min_{\sum_i n_i \le N} \sum_i F_i(n_i), \qquad F_i(n) = \left(\sum_{j>n} v_{i,j}^q \right)^{p/q}. \tag{9.1}$$

This is an allocation problem with unit costs. The relevant quantities are the successive reductions in each row's contribution to the error. Write $M_f = \sum_i m_i$ for the total number of coordinates.

Theorem 9.1 (Exact ordering rule for mixed approximation)**.** *Assume $0 < q \le p < \infty$ and put*

$$g_{i,n} = F_i(n-1) - F_i(n), \qquad 1 \le n \le m_i. \tag{9.2}$$

The gains are nonincreasing in every row. Let $g_1^* \ge \cdots \ge g_{M_f}^*$ be their decreasing arrangement, with ties ordered by their position within each row. For every integer N with $0 \le N \le M_f$,

$$\boxed{\sigma_N(f)^p = \sum_{k>N} g_k^*, \qquad 0 \le N \le M_f.} \tag{9.3}$$

Retaining the coefficients corresponding to the first N gains gives an optimal approximant. These supports are nested as N increases.

Proof. Fix a row and set $s = p/q \ge 1$ and $T_n = \sum_{j>n} v_{i,j}^q$. For $t, h \ge 0$, the function

$$\Delta_s(t,h) = (t+h)^s - t^s$$

is nondecreasing in both variables. The integral formula

$$\Delta_s(t,h) = s \int_0^h (t+x)^{s-1}\, dx$$

shows both monotonicities. Since $T_n \ge T_{n+1}$ and $v_{i,n}^q \ge v_{i,n+1}^q$,

$$g_{i,n} = \Delta_s(T_n, v_{i,n}^q) \ge \Delta_s(T_{n+1}, v_{i,n+1}^q) = g_{i,n+1}.$$

Apply Corollary 3.3 to (9.1). The terminal row errors vanish, giving (9.3). □

The proof uses the convexity of the scalar power applied to each row tail, so it also covers $p < 1$, when the mixed sequence space need not have a convex unit ball. For $p = q$, the gains are $v_{i,n}^p$, and the rule retains the largest coefficient magnitudes in the whole matrix.

Corollary 9.2 (Countable mixed sequence spaces)**.** *Let $0 < q \le p < \infty$ and let $f \in \ell^p(\ell^q)$ be a countable matrix. Define the row tails and gains as in (9.2), padding finite rows by zeros. Then, for every integer $N \ge 0$,*

$$\sigma_N(f)^p = \sum_{k>N} g_k^*.$$

There is one nested family of optimal coordinate supports for all N.

Proof. Each row belongs to ℓ^q. Its positive magnitudes can therefore be arranged in decreasing order, and its tail tends to zero. The calculation in Theorem 9.1 shows that the gains are nonincreasing within each row. Telescoping and summing nonnegative terms give

$$\sum_{i,n} g_{i,n} = \sum_i F_i(0) = \|f\|^p_{\ell^p(\ell^q)} < \infty.$$

Only finitely many positive gains exceed any given positive threshold, so the positive gains have a decreasing enumeration. Order ties by their position within each row, and append zeros if there are only finitely many positive gains. The first N gains form initial segments in the rows. Every N-term approximant can be replaced, without increasing the error, by a selection of this form. Its total gain is at most the sum of the first N gains, and choosing those gains attains that sum. The remaining error is the displayed tail, and the successive supports are nested. □

Example 9.3 (A smaller coefficient can give a larger improvement)**.** Take $p = 2$, $q = 1$, with rows $(1, 1)$ and $(3/2)$. The errors and marginal gains in the two rows are

$$F_1 = (4, 1, 0), \quad (g_{1,1}, g_{1,2}) = (3, 1), \qquad F_2 = (9/4, 0), \quad g_{2,1} = 9/4.$$

Ordering the gains as $3, 9/4, 1$ gives:

N	Retained coefficients	Residual error
0	none	$5/2$
1	one 1 from the first row	$\sqrt{13}/2$
2	that 1 and the $3/2$	1
3	all coefficients	0

For $N = 1$, retaining the largest coefficient, $3/2$, would leave error 2, which exceeds $\sqrt{13}/2$. The reduction in error depends on the whole row, not just on the magnitude of the retained coefficient.

Proposition 9.4 (Optimal supports need not be nested when $p < q$)**.** *For every $0 < p < q < \infty$, there is a finite matrix whose optimal one-term and two-term coordinate supports cannot be nested.*

Proof. Put $s = p/q < 1$, and choose $h \in (0, 1)$ with

$$2^s - 1 < h^p < 1.$$

Take rows $(1, 1)$ and (h). Retaining only h leaves the pth power of the residual equal to 2^s. Retaining one coefficient from the first row leaves $1 + h^p > 2^s$. Thus every optimal one-term support consists of the coefficient h.

Retaining both coefficients of the first row leaves the pth power of the residual equal to $h^p < 1$. Retaining h and one coefficient from the first row leaves pth power 1. Thus every optimal two-term support consists of the first row and omits h. The optimal supports cannot be nested. □

For example, $p = 1$, $q = 2$, $h = 1/2$ gives optimal residuals $\sqrt{2}$ and $1/2$ with one and two retained coefficients, respectively. The gains in the first row are $2^{p/q} - 1$ and 1, so they increase rather than decrease.

Summability of the initial row errors allows the variational representation to be applied for all $0 < p, q < \infty$.

Corollary 9.5 (A countable support model for mixed approximation)**.** *Let $0 < p, q < \infty$ and let $f \in \ell^p(\ell^q)$ be a countable matrix. There are a finite or countable set J and positive weights A_j, W_j, with $\sum_{j\in J} W_j < \infty$, such that*

$$D(B) := E(B, \mathbf{1}; \ell^0(A), \ell^1(W)) \le \sigma_{\lfloor B \rfloor}(f)^p \le 2D(B/2), \qquad B \ge 0. \tag{9.4}$$

The weights are independent of B. Every binary vector ξ with finite weighted support size gives a coordinate approximant h_ξ satisfying

$$\#\operatorname{supp} h_\xi \le 2\sum_{j\in J} A_j\xi_j, \qquad \|f - h_\xi\|^p_{\ell^p(\ell^q)} \le 2\sum_{j\in J} W_j(1-\xi_j). \tag{9.5}$$

For a finite set $Q \subset J$, put

$$\omega_Q = \sum_{j\notin Q} W_j, \tag{9.6}$$

$$D_Q(B) = \omega_Q + \min_{\substack{\xi\in\{0,1\}^Q \\ \sum_{j\in Q} A_j\xi_j \le B}} \sum_{j\in Q} W_j(1-\xi_j). \tag{9.7}$$

Then $0 \le D_Q(B) - D(B) \le \omega_Q$ for every $B \ge 0$, and

$$\max\{0, D_Q(N) - \omega_Q\} \le \sigma_N(f)^p \le 2D_Q(N/2), \qquad N \in \mathbb{N}_0. \tag{9.8}$$

Applying the reconstruction map to a minimizer in (9.7) with bound $N/2$ gives an approximant $h_{N,Q}$ with

$$\#\operatorname{supp} h_{N,Q} \le N, \qquad \|f - h_{N,Q}\|^p_{\ell^p(\ell^q)} \le 2D_Q(N/2).$$

For any increasing exhaustion $Q_m \uparrow J$, the optimal error functions D_{Q_m} of the finite models converge to D uniformly for $B \ge 0$.

Proof. Arrange the positive magnitudes in each row in decreasing order, padding finite rows by zeros, and define

$$F_i(n) = \left(\sum_{j>n} v_{i,j}^q\right)^{p/q}, \qquad C_i(n) = n, \qquad n \in \mathbb{N}_0.$$

For any fixed coordinate support, reproducing the selected coefficients minimizes the residual. For a prescribed number of retained coefficients in a row, choosing its largest magnitudes minimizes that row's contribution. Both assertions follow from monotonicity of the positive scalar powers. The minimum for the resulting component allocation with total cost at most B is therefore exactly $\sigma_{\lfloor B \rfloor}(f)^p$.

The upward attainable diagram of the ith row is

$$\Gamma_i = \bigcup_{n\ge 0} [n, \infty) \times [F_i(n), \infty).$$

It is closed because only finitely many levels n can occur under a fixed bound on C_i. Moreover,

$$F_i(n) \longrightarrow 0, \qquad \sum_i F_i(0) = \|f\|^p_{\ell^p(\ell^q)} < \infty.$$

Theorem 7.4 therefore applies with $a = b = 2$ and proves (9.4). Its reconstruction map associates to each binary vector of finite weighted support size a sequence of original row levels n_i such that

$$\sum_i n_i \le 2 \sum_j A_j \xi_j, \qquad \sum_i F_i(n_i) \le 2 \sum_j W_j (1 - \xi_j).$$

The first sum is finite and its terms are nonnegative integers. Thus only finitely many positive row levels occur. Retaining the first n_i coefficients in each row gives a finitely supported coordinate approximant and proves (9.5).

The quantity $D_Q(B)$ is the minimum of the full model error over binary vectors that vanish outside Q. This restriction gives $D(B) \le D_Q(B)$. Conversely, truncating any admissible binary vector to Q does not increase its cost and increases its error by

$$\sum_{j \notin Q} W_j \xi_j \le \omega_Q.$$

Taking the infimum over the admissible binary vectors gives $D_Q(B) \le D(B) + \omega_Q$. Combining this estimate with (9.4) proves (9.8). The finite minimum defining $D_Q(N/2)$ is attained. Extend a minimizer by zero outside Q and apply (9.5); this yields the stated approximant $h_{N,Q}$. Finally, $\omega_{Q_m} \to 0$ by summability, and the estimate independent of B proves uniform convergence of the optimal error functions. □

For finite matrices, Corollary 11.2 constructs a nested family of coordinate approximants using the minimizers of penalized problems from Section 11.

10 Penalized values and K–J representations

We now return to finitely many components, each with finitely many admissible choices. The geometric discretization estimates the optimal error under the constraint $C(z) \le B$. To represent the penalized minimum exactly, we group consecutive increments along the lower convex hull of each row. Selecting whole blocks gives vertices of this hull, so the penalized minima are attained by original admissible choices.

10.1 Convex hull blocks

Take normalized Pareto chains as in (4.1). In row i, let

$$0 = j_{i,0} < j_{i,1} < \cdots < j_{i,L_i} = N_i$$

be the vertices of the lower convex hull of the points $(c_{i,n}, e_{i,n})$. Adjacent collinear edges may be merged. For a hull block $\beta = (i, \ell)$ set

$$c_\beta = c_{i,j_{i,\ell}} - c_{i,j_{i,\ell-1}}, \qquad d_\beta = e_{i,j_{i,\ell-1}} - e_{i,j_{i,\ell}}, \qquad r_\beta = c_\beta / d_\beta. \tag{10.1}$$

All block weights are positive. The lower convex boundary has nondecreasing slopes $-d_\beta/c_\beta$, so the ratios $r_{i,\ell}$ are nondecreasing within each row. They are strictly increasing after adjacent collinear edges are merged.

Theorem 10.1 (Exact positive representation of the penalized value)**.** *For all $\lambda > 0$,*

$$\mathcal{K}(\lambda) = \sum_{\beta} \min(c_\beta, \lambda d_\beta). \tag{10.2}$$

Selecting the blocks with $r_\beta \le \lambda$ gives a minimizer of the penalized problem. The selected blocks form an initial segment in each row and end at a hull vertex corresponding to an original admissible choice. Selecting all blocks with $r_\beta < \lambda$ is also optimal; at a tied ratio any initial segment of the tied blocks is allowed.

Proof. Rows are independent, so $\mathcal{K}$ is the sum of the row minima. A linear functional has the same minimum on a finite set and its convex hull. For the functional $(c, e) \mapsto c + \lambda e$, that minimum can be taken at a lower hull vertex. At vertex $j_{i,\ell}$ the value is

$$c_{i,j_{i,\ell}} + \lambda e_{i,j_{i,\ell}} = \sum_{h=1}^{\ell} c_{(i,h)} + \lambda \sum_{h=\ell+1}^{L_i} d_{(i,h)},$$

where $c_{(i,h)}$ and $d_{(i,h)}$ are the block weights in (10.1). Moving from vertex $j_{i,\ell-1}$ to vertex $j_{i,\ell}$ changes this expression by

$$c_{(i,\ell)} - \lambda d_{(i,\ell)} = d_{(i,\ell)}(r_{(i,\ell)} - \lambda).$$

As ℓ increases, the sign can change only from negative to positive. An optimal initial segment therefore selects the blocks for which $c_\beta \le \lambda d_\beta$. This proves the formula and describes the minimizers. □

Selecting individual increments instead of whole blocks can change the minimum. For example, a row with states $(0, 2), (3, 1), (4, 0)$ has

$$\mathcal{K}(\lambda) = \min(2\lambda, 4).$$

Selecting its individual increments independently would give $\min(3, \lambda) + \min(1, \lambda)$, which is 3 at $\lambda = 2$ instead of the correct value 4. The smaller expression selects the second increment without the first. The hull representation groups both increments into one admissible block with weights 4 and 2.

10.2 Exact representations and geometric discretization

Let Ω be the finite set of original row increments. Give (i, n) the weights

$$a_{i,n} = c_{i,n} - c_{i,n-1}, \qquad w_{i,n} = e_{i,n-1} - e_{i,n}.$$

For vectors on Ω define

$$A(v) = \sum_{j : v_j \neq 0} a_j, \qquad W(v) = \sum_j w_j |v_j|, \qquad \mathcal{J}(t, v) = \max\{A(v), tW(v)\}.$$

The indicator v_β of a hull block satisfies $A(v_\beta) = c_\beta$ and $W(v_\beta) = d_\beta$. These indicators are disjoint and sum to $\mathbf{1}_\Omega$. Consider the subgroup of binary vectors that are constant on each hull block. With the two inherited group norms, this is the pair of weighted groups indexed by the blocks, with weights c_β, d_β. By Theorem 10.1, its K-functional at $\mathbf{1}$ equals the original penalized value. Every term of the representations below belongs to this subgroup.

For each distinct block ratio r, let $u_r = \sum_{r_\beta = r} v_\beta$. Then $A(u_r) = rW(u_r)$, and Theorem 10.1 immediately gives

$$\mathbf{1}_\Omega = \sum_r u_r, \qquad \boxed{\mathcal{K}(\lambda) = \sum_r \min(1, \lambda/r)\mathcal{J}(r, u_r).} \tag{10.3}$$

The parameter values in this exact formula are determined by the data.

Theorem 10.2 (Representation on a fixed geometric scale)**.** *Fix $\rho > 1$, put $q_0 = \sqrt{\rho}$, and let $s_k = \rho^k$, $k \in \mathbb{Z}$. Group whole hull blocks by*

$$s_k/q_0 < r_\beta \le q_0 s_k, \qquad u_k = \sum_{s_k/q_0 < r_\beta \le q_0 s_k} v_\beta.$$

There are finitely many nonzero u_k, they are disjoint and binary, and they sum to $\mathbf{1}_\Omega$. For every $\lambda > 0$,

$$\boxed{\mathcal{K}(\lambda) \le \sum_k \min(1, \lambda/s_k)\mathcal{J}(s_k, u_k) \le \sqrt{\rho}\,\mathcal{K}(\lambda).} \tag{10.4}$$

Moreover, selecting the groups with $s_k \le \lambda$ gives an original admissible vector with penalized value at most $\sqrt{\rho}\,\mathcal{K}(\lambda)$. These vectors form a nested family.

Proof. The group indices are nondecreasing within each row because the block ratios increase. Selecting all groups with $s_k \le \lambda$ therefore selects an initial segment of whole blocks in every row and gives an original admissible choice. Its penalized value is

$$\sum_{s_k \le \lambda} A(u_k) + \lambda \sum_{s_k > \lambda} W(u_k),$$

which is at most the middle expression in (10.4). Taking the minimum over admissible original choices gives the lower bound.

For a block in group k, the bounds on its ratio imply

$$\max(c_\beta, s_k d_\beta) \le q_0 c_\beta, \qquad \max(c_\beta, s_k d_\beta) \le q_0 s_k d_\beta.$$

After multiplication by $\min(1, \lambda/s_k)$, the expression is bounded both by $q_0 c_\beta$ and by $q_0 \lambda d_\beta$. Hence

$$\min(1, \lambda/s_k)\max(c_\beta, s_k d_\beta) \le q_0 \min(c_\beta, \lambda d_\beta).$$

Because the blocks are disjoint,

$$\mathcal{J}(s_k, u_k) = \max\left(\sum_{\beta \in k} c_\beta, s_k \sum_{\beta \in k} d_\beta\right) \le \sum_{\beta \in k} \max(c_\beta, s_k d_\beta).$$

Summing and using (10.2) proves the upper bound. The same estimates bound the penalized value of the vector obtained by selecting these groups. □

Proposition 10.3 (Sharpness for binary decompositions into blocks)**.** *For the fixed grid $s_k = \rho^k$, the constant $\sqrt{\rho}$ in (10.4) cannot be decreased uniformly for disjoint binary decompositions into whole hull blocks.*

Proof. Take one block with weights $q_0 = \sqrt{\rho}$ and 1. A disjoint binary representation of its indicator assigns it to exactly one parameter value s_k. For this assignment,

$$\sup_{\lambda>0} \frac{\min(1, \lambda/s_k)\max(q_0, s_k)}{\min(q_0, \lambda)} = \max(q_0/s_k, s_k/q_0).$$

For $s_k = \rho^k$ the smallest possible value is q_0, attained at $k = 0$ and $k = 1$. □

For dyadic parameter values, the optimal constant for binary representations by whole blocks is $\sqrt{2}$. The same decomposition estimates the K-functional at every parameter. We next obtain a continuous representation and relate it to the theorem of Cwikel–Jawerth–Milman for general Banach pairs [3].

10.3 A continuous model

Let $\mathcal{B}$ be the finite set of hull blocks and consider the Banach pair

$$\mathbf{B} = (B_0, B_1) = (\ell^1(\mathcal{B}, c), \ell^1(\mathcal{B}, d)), \qquad \|x\|_{B_0} = \sum_{\beta\in\mathcal{B}} c_\beta |x_\beta|, \quad \|x\|_{B_1} = \sum_{\beta\in\mathcal{B}} d_\beta |x_\beta|.$$

The block weights are strictly positive. Writing $e = \mathbf{1}_{\mathcal{B}}$, coordinatewise minimization and Theorem 10.1 give

$$K(\lambda, e; \mathbf{B}) = \sum_{\beta\in\mathcal{B}} \min(c_\beta, \lambda d_\beta) = \mathcal{K}(\lambda). \tag{10.5}$$

Indeed, the minimum of $c_\beta|z| + \lambda d_\beta|1 - z|$ over scalars z is $\min(c_\beta, \lambda d_\beta)$. This Banach pair has the same penalized minima and permits Bochner integration. The selection rule in Theorem 10.2 still gives original admissible vectors by selecting whole blocks.

Put

$$J_{\mathbf{B}}(s, x) = \max\{\|x\|_{B_0}, s\|x\|_{B_1}\}.$$

Let $\widehat{u}_k$ denote u_k from Theorem 10.2 written in block coordinates: its βth coordinate is one when $s_k/\sqrt{\rho} < r_\beta \le \sqrt{\rho}\, s_k$, and zero otherwise. On these vectors the endpoint norms equal $A(u_k)$ and $W(u_k)$. Hence (10.4) holds with $J_{\mathbf{B}}(s_k, \widehat{u}_k)$ in place of $\mathcal{J}(s_k, u_k)$.

Corollary 10.4 (Continuous representation of the discrete model)**.** *For every $\rho > 1$ there is a simple $B_0 \cap B_1$-valued function U_ρ, with compact support in $(0, \infty)$, such that*

$$e = \int_0^\infty U_\rho(s)\, \frac{ds}{s}. \tag{10.6}$$

Define

$$\mathcal{D}_\rho(\lambda) = \int_0^\infty \min\{1, \lambda/s\} J_{\mathbf{B}}(s, U_\rho(s))\, \frac{ds}{s}. \tag{10.7}$$

The same representation satisfies, for every $\lambda > 0$,

$$\mathcal{K}(\lambda) \le \mathcal{D}_\rho(\lambda) \le \rho\, \mathcal{K}(\lambda), \tag{10.8}$$

$$\int_0^\lambda \|U_\rho(s)\|_{B_0}\, \frac{ds}{s} + \lambda \int_\lambda^\infty \|U_\rho(s)\|_{B_1}\, \frac{ds}{s} \le \mathcal{D}_\rho(\lambda). \tag{10.9}$$

It also satisfies the pointwise bound

$$J_{\mathbf{B}}(s, U_\rho(s)) \le \frac{\rho}{\log \rho}\, \mathcal{K}(s), \qquad s > 0. \tag{10.10}$$

Proof. Set $s_k = \rho^k$ and use the disjoint logarithmic intervals

$$I_k = [s_k \rho^{-1/2}, s_k \rho^{1/2}), \qquad k \in \mathbb{Z}.$$

They cover $(0, \infty)$ and each has ds/s-measure $\log \rho$. Define

$$U_\rho(s) = \frac{\widehat{u}_k}{\log \rho}, \qquad s \in I_k. \tag{10.11}$$

Only finitely many $\widehat{u}_k$ are nonzero, and their sum is e. Thus (10.6) is an absolutely convergent Bochner integral.

For $s \in I_k$, homogeneity of the endpoint norms gives

$$\begin{aligned} J_{\mathbf{B}}(s, \widehat{u}_k) &\le \max\{1, s/s_k\} J_{\mathbf{B}}(s_k, \widehat{u}_k), \\ \min\{1, \lambda/s\} &\le \max\{1, s_k/s\} \min\{1, \lambda/s_k\}. \end{aligned}$$

The product of the two comparison factors is at most $\sqrt{\rho}$. Integrating on I_k and summing, we therefore obtain

$$\mathcal{D}_\rho(\lambda) \le \sqrt{\rho} \sum_k \min\{1, \lambda/s_k\} J_{\mathbf{B}}(s_k, \widehat{u}_k) \le \rho \, \mathcal{K}(\lambda),$$

where the last step is Theorem 10.2.

Splitting (10.6) at $s = \lambda$ gives a decomposition of e into a B_0 part and a B_1 part. The triangle inequality and the definition of $J_{\mathbf{B}}$ show that

$$\begin{aligned} K(\lambda, e; \mathbf{B}) &\le \int_0^\lambda \|U_\rho(s)\|_{B_0} \frac{ds}{s} + \lambda \int_\lambda^\infty \|U_\rho(s)\|_{B_1} \frac{ds}{s} \\ &\le \mathcal{D}_\rho(\lambda). \end{aligned}$$

Together with (10.5), this proves (10.8) and (10.9).

For the pointwise estimate, the kth summand of (10.4), evaluated at $\lambda = s_k$, gives

$$J_{\mathbf{B}}(s_k, \widehat{u}_k) \le \sqrt{\rho} \, \mathcal{K}(s_k).$$

The monotonicity of $\mathcal{K}(t)$ and $\mathcal{K}(t)/t$ gives $\mathcal{K}(s_k) \le \max\{1, s_k/s\} \mathcal{K}(s)$. Combining this with the first inequality above and (10.11) proves (10.10). □

As $\rho \downarrow 1$, the integral comparison factor ρ approaches one, whereas the pointwise factor $\rho/\log \rho$ diverges. Thus the integral estimate improves as the bound for the density increases.

The two parts of the integral give the Gagliardo coordinates. Put

$$A_\rho(\lambda) = \int_0^\lambda J_{\mathbf{B}}(s, U_\rho(s)) \frac{ds}{s}, \tag{10.12}$$

$$B_\rho(\lambda) = \int_\lambda^\infty J_{\mathbf{B}}(s, U_\rho(s)) \frac{ds}{s^2}. \tag{10.13}$$

Then

$$\mathcal{D}_\rho(\lambda) = A_\rho(\lambda) + \lambda B_\rho(\lambda), \qquad \mathcal{D}'_\rho(\lambda) = B_\rho(\lambda), \qquad \lambda > 0. \tag{10.14}$$

The derivative follows by differentiating (10.7) under the integral; the function U_ρ is simple and supported away from zero and infinity. Thus A_ρ and B_ρ are the Gagliardo coordinates of the concave function $\mathcal{D}_\rho$. The two vectors

$$e_{0,\lambda} = \int_0^\lambda U_\rho(s) \frac{ds}{s}, \qquad e_{1,\lambda} = \int_\lambda^\infty U_\rho(s) \frac{ds}{s}$$

satisfy $e = e_{0,\lambda} + e_{1,\lambda}$ and

$$\|e_{0,\lambda}\|_{B_0} \le A_\rho(\lambda), \qquad \|e_{1,\lambda}\|_{B_1} \le B_\rho(\lambda).$$

The general Banach-pair setting. The passage from sums to integrals in [3, Theorems 2.13 and 3.1, pp. 77–79] groups vectors according to the ratios of their endpoint norms and assigns each group to a logarithmic interval. For a Banach pair $\mathbf{A} = (A_0, A_1)$, suppose $f = \sum_k v_k$, with $v_k \in A_0 \cap A_1$ and $s_k = \rho^k$, satisfies

$$\sum_k \min\{1, \lambda/s_k\} J_{\mathbf{A}}(s_k, v_k) \le CK(\lambda, f; \mathbf{A}), \qquad \lambda > 0.$$

At $\lambda = 1$ this gives

$$\sum_k \|v_k\|_{A_0+A_1} \le \sum_k \min\{1, 1/s_k\} J_{\mathbf{A}}(s_k, v_k) < \infty,$$

where the sum norm is $K(1, \cdot; \mathbf{A})$. Thus $U(s) = v_k/\log\rho$ on I_k represents f by an absolutely convergent Bochner integral. The preceding proof gives the integral K–J bound with constant $C\sqrt{\rho}$ and the separate endpoint estimates.

The strong fundamental lemma supplies the discrete representation. Its continuous form in [3, Theorem 3.1] assumes a mutually closed pair, meaning that each endpoint agrees isometrically with its Gagliardo completion in the sum space, and

$$f \in \Sigma^\circ(\mathbf{A}) := \overline{A_0 \cap A_1}^{A_0+A_1}.$$

For every $\varepsilon > 0$ it gives a single $A_0 \cap A_1$-valued function V_ε, strongly measurable on compact subintervals and absolutely Bochner integrable in $A_0 + A_1$, with

$$\begin{aligned} f &= \int_0^\infty V_\varepsilon(s)\,\frac{ds}{s}, \\ \int_0^\infty \min\{1, \lambda/s\} J_{\mathbf{A}}(s, V_\varepsilon(s))\,\frac{ds}{s} &\le (1+\varepsilon)^2(3+2\sqrt{2})K(\lambda, f; \mathbf{A}), \qquad \lambda > 0. \end{aligned} \tag{10.15}$$

Corollary 10.4 applies this passage to the weighted model of Theorem 10.2, whose positive block structure gives the bound $\rho\,\mathcal{K}(\lambda)$.

10.4 Constrained minimizers and positive linear penalties

The family of penalized minimizers need not contain every constrained minimizer. Consider two independent binary coordinates with weights

$$(c_1, d_1) = (2, 3), \qquad (c_2, d_2) = (4, 4).$$

Under the constraint $C(z) \le 2$, the unique minimizer selects the first coordinate; under $C(z) \le 4$, it selects the second. These supports are not nested. The penalized value is

$$\mathcal{K}(\lambda) = \min(2, 3\lambda) + \min(4, 4\lambda).$$

As the ratio threshold increases, it selects neither coordinate, then the first, then both. Selecting the second coordinate alone minimizes the error under $C(z) \le 4$ but never minimizes a positive linear penalty.

The same distinction appears in the minimum values. For the problem with one binary coordinate in Example 3.2,

$$F(B) = \begin{cases} 1, & 0 \le B < 2, \\ 0, & B \ge 2, \end{cases} \qquad \mathcal{K}(\lambda) = \min(2, \lambda),$$

but

$$\sup_{\lambda > 0} \frac{\mathcal{K}(\lambda) - B}{\lambda} = (1 - B/2)_+.$$

The supporting-line transform gives the lower convex envelope of the optimal error function. It does not recover the original step function. The geometric representation estimates the original optimal error function directly, with the stated factors for C and E. Section 12 describes recovery by a family of power penalties.

11 A nested family of admissible approximate solutions

The exact representation of the penalized minima gives a family of admissible allocations for every bound $B \ge 0$. The estimate below is the finite separable form of the weighted-sum principle of Halffmann, Ruzika, Thielen, and Willems [8]. Ordering the convex hull blocks determines the successive choices in this family.

Sort all L hull blocks from Section 10 by nondecreasing ratio $r_\beta = c_\beta / d_\beta$, preserving the order of tied blocks within each row. Relabel them as $(c_1, d_1), \dots, (c_L, d_L)$. Selecting the first k blocks selects an initial segment in every row and therefore gives an original admissible allocation z^k. Write

$$C_k = \sum_{j \le k} c_j, \qquad E_k = \sum_{j > k} d_j, \qquad 0 \le k \le L. \tag{11.1}$$

These allocations form a nested family in the ordered levels of each component. The optimal error function for this family is

$$H(B) = \min\{E_k : C_k \le B\}.$$

Theorem 11.1 (Approximation by penalized minimizers)**.** *For every $\alpha > 1$ and every $B \ge 0$,*

$$\boxed{F(B) \le H(B) \le \frac{\alpha}{\alpha - 1} F(B/\alpha).} \tag{11.2}$$

Equivalently, the largest initial segment with $C_k \le \alpha B$ satisfies

$$C(z^k) \le \alpha B, \qquad E(z^k) \le \frac{\alpha}{\alpha - 1} F(B).$$

The same family gives these estimates for every $B \ge 0$ and every $\alpha > 1$.

Proof. Each z^k is admissible in the original problem, so $F(B) \le H(B)$. It remains to prove the equivalent upper bound

$$H(\alpha B) \le \frac{\alpha}{\alpha - 1} F(B).$$

If $B \ge C_L$, the terminal choice has zero error and the assertion follows. Otherwise choose j so that $C_j \le B < C_{j+1}$, and put

$$\theta = \frac{B - C_j}{C_{j+1} - C_j}, \qquad r_* = \frac{c_{j+1}}{d_{j+1}}.$$

Both adjacent initial segments minimize $C + r_* E$, by Theorem 10.1. This remains true at tied block ratios. Therefore

$$\mathcal{K}(r_*) = C_j + r_* E_j = C_{j+1} + r_* E_{j+1}.$$

Every original admissible vector with $C(z) \le B$ has error at least

$$\frac{\mathcal{K}(r_*) - B}{r_*} = (1-\theta)E_j + \theta E_{j+1} =: \ell(B). \tag{11.3}$$

It follows that $\ell(B) \le F(B)$.

Choose the largest initial segment with $C_k \le \alpha B$. If $C_{j+1} \le \alpha B$, its error is at most $E_{j+1} \le \ell(B) \le F(B)$. Otherwise $C_{j+1} > \alpha B$. Since $C_j \ge 0$,

$$\theta = \frac{B - C_j}{C_{j+1} - C_j} \le \frac{B}{C_{j+1}} < \frac{1}{\alpha}.$$

Hence

$$\ell(B) \ge (1-\theta)E_j \ge \left(1 - \frac{1}{\alpha}\right) E_j.$$

The selected initial segment contains at least the first j blocks, so its error is at most

$$E_j \le \frac{\alpha}{\alpha - 1}\ell(B) \le \frac{\alpha}{\alpha - 1}F(B).$$

This proves the estimate, including $B = 0$. □

11.1 Construction and evaluation for a prescribed bound

Suppose the admissible component choices are explicitly given, and let $M = \sum_i |\mathcal{A}_i|$. Construct the family as follows.

1. Sort the pairs $(C_i(a), E_i(a))$ in each component by their first coordinate and remove dominated pairs. Retain an admissible choice corresponding to each remaining pair.
2. Compute the lower convex hull of each component using a stack. Process the points in increasing order of the first coordinate, removing a middle vertex whenever the two consecutive slopes fail to increase. Each vertex is added to and removed from the stack at most once, so this step takes linear time after sorting.
3. Form the hull blocks and order their ratios, preserving row order when ratios are equal. For each block, store the partial sum C_k and residual error E_k, together with the corresponding change of component choice.

These steps require $O(M \log M)$ arithmetic and comparison operations. The list has at most M blocks and determines the entire family without enumerating the product of the admissible sets. For a prescribed bound B, binary search identifies the appropriate initial segment. The corresponding component choices are obtained from the stored changes, or from the last selected block in each row.

For unnormalized data, let C_0, R_0 be the minima of C and E from Section 6.1. For $B \ge C_0$, apply the normalized construction with bound $B - C_0$. The resulting vector satisfies

$$C(z^k) \le C_0 + \alpha(B - C_0),$$
$$E(z^k) \le R_0 + \frac{\alpha}{\alpha - 1}\big(F(B) - R_0\big).$$

Thus the factors apply to C and E after their respective minima have been subtracted.

11.2 Nested approximation in finite mixed sequence spaces

For finite matrices, the same construction gives a nested family of coordinate approximations.

Corollary 11.2 (Nested coordinate approximations for finite matrices)**.** *Let f be a finite matrix and let $0 < p, q < \infty$. There is a nested family of coordinate supports such that, for every integer $N \geq 0$ and every $\alpha > 1$, the family contains an approximant h satisfying*

$$\#\operatorname{supp} h \leq \lfloor \alpha N \rfloor, \qquad \|f - h\|_{\ell^p(\ell^q)} \leq \left(\frac{\alpha}{\alpha - 1}\right)^{1/p} \sigma_N(f).$$

Compute the finite row tails, take their lower convex hulls, and order the resulting block ratios to construct the family.

Proof. Apply Theorem 11.1 to the row pairs $(n, F_i(n))$. Each selected hull vertex corresponds to retaining an initial segment of the row's decreasing magnitudes. The corresponding coordinate supports are therefore nested. Here C is the integer number of retained coordinates, and E is the pth power of the error in the mixed norm. Taking pth roots gives the stated inequalities. □

11.3 Allocation among hierarchical corrections

A hierarchical decomposition may provide corrections at several scales. For each correction, one can choose whether to retain it and, if so, at which of finitely many accuracies to compute it. Additive upper bounds for the norm of the approximant and its residual error allow us to apply the finite allocation theorem.

Proposition 11.3 (Approximation from finite sets of corrections)**.** *Let $\mathcal{B}$ and Y be Banach spaces, and let T be a linear operator on $\mathcal{B}$ with values in a vector space containing Y. Suppose $f = \sum_{i=1}^{M} g_i + r_M$ in Y. For each i, let $\mathcal{A}_i$ be a nonempty finite set of choices, represented by $u_{i,a} \in \mathcal{B}$ with $Tu_{i,a} \in Y$. Set*

$$C_i(a) = \|u_{i,a}\|_{\mathcal{B}}, \qquad E_i(a) = \|g_i - Tu_{i,a}\|_Y.$$

Every choice $a = (a_1, \ldots, a_M)$ gives $U_a = \sum_i u_{i,a_i}$ satisfying

$$\begin{aligned} \|U_a\|_{\mathcal{B}} &\leq \sum_i C_i(a_i), \\ \|f - TU_a\|_Y &\leq \|r_M\|_Y + \sum_i E_i(a_i). \end{aligned} \tag{11.4}$$

Theorem 1.1 therefore applies to these additive upper bounds with its stated factors for the two criteria, using Section 6.1 for nonzero component minima.

Proof. Both bounds follow from the triangle inequality and the identity $f - TU_a = r_M + \sum_i (g_i - Tu_{i,a_i})$. The choices in the different components are independent, so the sums of C_i and E_i have the required form. □

For binary retention of corrections w_i, take $g_i = Tw_i$ and the choice set $\{0, w_i\}$. A selected set S then satisfies

$$\left\| \sum_{i \in S} w_i \right\|_{\mathcal{B}} \le \sum_{i \in S} \|w_i\|_{\mathcal{B}},$$

$$\left\| f - T \sum_{i \in S} w_i \right\|_{Y} \le \|r_M\|_Y + \sum_{i \notin S} \|Tw_i\|_Y.$$

Minimizing the latter upper bound subject to a bound on $\sum_{i \in S} \|w_i\|_{\mathcal{B}}$ is a knapsack problem. Theorem 11.1 also supplies a family of choices minimizing positive linear penalties, valid for every bound B. Cancellation may make the actual residual smaller than this upper bound.

For the critical divergence problem in dimension $d \ge 2$, one may take $T = \operatorname{div}$, $Y = L^d_{\#}(\mathbb{T}^d)$, and $\mathcal{B} = L^\infty$ or $L^\infty \cap W^{1,d}$, where $L^d_{\#}$ denotes mean-zero data and the solution spaces are vector-valued, with their corresponding norms. The hierarchical construction supplies the fields and their estimates [16, 2]; the finite minimization problem selects among them. The companion Mosco analysis [11] studies stability of selected representations as the admissible approximation spaces change.

11.4 Necessary approximation factors for penalized minimizers

The relation between the factors for C and E has a geometric explanation. An admissible choice that minimizes no positive linear penalty may give a better constrained approximation than every penalized minimizer.

Proposition 11.4 (Necessary relation between approximation factors)**.** *Let $\alpha, \beta > 1$. Using only minimizers of positive linear penalties, the estimates*

$$C(z) \le \alpha B, \qquad E(z) \le \beta F(B)$$

can hold uniformly for every finite set of pairs $(C(z), E(z))$ only if $\alpha^{-1} + \beta^{-1} \le 1$.

Proof. If $\alpha^{-1} + \beta^{-1} > 1$, choose numbers

$$\alpha < T < \frac{\beta}{\beta - 1}, \qquad 1 - \frac{1}{T} < e < \frac{1}{\beta}.$$

Consider the three pairs $(0,1), (1,e), (T,0)$. The middle pair lies strictly above the chord joining the endpoints, so it minimizes no positive linear penalty. At $B = 1$, it does minimize the error under $C(z) \le B$. The penalized minimizers are the endpoints: the choice with $C(z) = 0$ has error $1 > \beta e$, and the choice with $E(z) = 0$ has $C(z) = T > \alpha B$. Neither satisfies both bounds. □

Theorem 11.1 attains the boundary value $\beta = \alpha/(\alpha - 1)$ in the separable problem. The necessary condition applies to methods restricted to penalized minimizers. Methods that use other admissible choices are not subject to it. For approximation by positive linear penalties in more general settings, see [1].

12 Recovering the optimal error function from power penalties

Let $\mathcal{A}$ be a nonempty admissible set, with finite nonnegative functions C, E, and suppose $C(z) = 0$ for some $z \in \mathcal{A}$. Define

$$F(B) = \inf_{C(z) \le B} E(z), \qquad B \ge 0.$$

Then F is finite and nonincreasing. The set $\mathcal{A}$ may be the subset of a product of component choices on which both total criteria are finite; in that case E denotes the total error. For $p > 0$ define

$$\mathcal{K}_p(\lambda) = \inf_{z\in\mathcal{A}} (C(z) + \lambda E(z)^p), \qquad H_p(B) = \sup_{\lambda>0} \frac{\mathcal{K}_p(\lambda) - B}{\lambda}. \tag{12.1}$$

Theorem 12.1 (Recovery from power penalties)**.** *For every $B \geq 0$, $\delta > 0$, and $p > 0$,*

$$\frac{\delta}{B+\delta} F(B+\delta)^p \leq H_p(B) \leq F(B)^p. \tag{12.2}$$

Consequently, at every $B \geq 0$ at which F is right-continuous,

$$\lim_{p\to\infty} H_p(B)^{1/p} = F(B). \tag{12.3}$$

This includes every $B \geq 0$ for finite choice sets and for the countable problems of Theorem 7.4.

Proof. For any competitor with $C(z) \leq B$,

$$\frac{\mathcal{K}_p(\lambda) - B}{\lambda} \leq E(z)^p + \frac{C(z) - B}{\lambda} \leq E(z)^p.$$

Taking the infimum over such competitors proves the upper bound. Also $H_p(B) \geq 0$, since $\mathcal{K}_p(\lambda) \geq 0$ and $-B/\lambda \to 0$ as $\lambda \to \infty$.

Set $A = F(B+\delta)^p$. If $A = 0$, the lower bound is already proved. Otherwise put $\alpha = A/(B+\delta) > 0$. For any z, either $C(z) \leq B+\delta$, in which case $E(z)^p \geq A$, or $C(z) > B+\delta$, in which case $\alpha C(z) > A$. Nonnegativity then gives

$$E(z)^p + \alpha C(z) \geq A \quad \text{for every } z.$$

Taking $\lambda = 1/\alpha$ in (12.1) yields

$$H_p(B) \geq \inf_z \{E(z)^p + \alpha(C(z) - B)\} \geq A - \alpha B = \frac{\delta}{B+\delta} A.$$

This proves (12.2). Take pth roots and let $p \to \infty$. The lower limit is at least $F(B+\delta)$ and the upper limit is at most $F(B)$. Right-continuity at B permits $\delta \downarrow 0$ and proves (12.3). The optimal error functions of finite choice sets are right-continuous, and Theorem 7.4 proves the same property for the countable problems considered here. □

This argument with power penalties was developed in the unpublished Paseky notes [9, Section 3.5, Lemma 11]. The estimate (12.2) gives an explicit recovery statement using the whole family $\mathcal{K}_p$. For a sum of row errors, applying the power to that total generally destroys separability:

$$\left(\sum_i E_i(z_i)\right)^p \neq \sum_i E_i(z_i)^p.$$

The formula recovers the optimal error function at points of right-continuity. Efficient computation of the power penalties requires further structure when taking powers destroys separability.

13 Further directions

Section 7 treats infinitely many choices and countably many components for which the minimum errors under $C_i = 0$ are summable. Its finite approximations are controlled by the omitted error weights. A further problem is to develop analogous maps when the attainable diagrams are not closed, with a precise distinction between points in the closure and values attained by admissible choices. One may then need to allow an additive error and a positive increase in the bound B. For diagrams with infinite error at $C = 0$, the two-sided construction of the unpublished notes [9, Section 4.2] provides a starting point. Extending it to countable families with explicit attainment and finite approximation bounds requires appropriate summability conditions.

Independence of the components is used throughout. For coupled constraints, the question is whether a replacement can preserve admissibility without increasing either criterion. Such a construction would replace the independent operations on initial segments used here.

Theorem 5.2 determines the sharp factor for C_d in replacement by initial segments without increasing model error for every $a > 1$ and $b \geq 1$ under its row hypotheses. For $a, b > 1$, combining it with the geometric maps gives the factors for C and E for the original problem stated in Corollary 7.7. A further quantitative problem is to determine the best simultaneous factors for C and E when the weights and reconstruction maps may be chosen freely. Allowing a prescribed increase in model error during replacement by initial segments gives another set of possible factors. The obstruction in Proposition 11.4 gives a necessary condition when the reconstructed choices must minimize positive linear penalties.

The mixed-sequence application raises a related question about nonlinear approximation. Exact nested optimal supports exist throughout $0 < q \leq p < \infty$, while for $p < q$ the obstruction already occurs in a three-coordinate example. Structural conditions on the data, or on more general lattices, may permit exact nesting or improve the constants for a nested family of approximate solutions.

A Variational problems of Micchelli and Pinkus

We apply the representation theorem to three variational problems considered by Micchelli and Pinkus [10]: diagonal approximation, smoothing, and linear recovery from noisy data. Their work uses Gagliardo diagrams to relate constrained and penalized problems and to describe their solutions. Here we construct weighted interpolation models and reconstruction maps from binary vectors to admissible approximations. The separate estimates for C and E hold for every bound B and penalty parameter. Depending on the problem, the resulting approximation is a sequence, a family of fitted functions, or a bounded linear recovery operator.

The diagonal problem allows countably many coordinates and positive exponents below one. The smoothing problem imposes a bound on the sum of squared derivative seminorms over countably many components. Finitely supported binary vectors give admissible solutions and bounds for the full problem. Once the selected points on the component boundaries and their representatives are known, finite subsets of the weights define finite constrained minimization problems. The omitted weights bound the truncation error uniformly in the parameter. Examples A.5 and A.7 compute these weights and representatives explicitly. In the smoothing example, the same finite family of fits is used for every bound B.

A.1 Diagonal approximation and finite error bounds

Micchelli and Pinkus [10, equation (4.4)] consider the balance between $\|f - Dg\|_p$ and $\|g\|_q$ for $f, g \in \mathbb{R}^n$, a finite diagonal matrix D, and $p, q \in [1, \infty]$. Theorem 7.4 also applies to finite and countable diagonal systems for $0 < p, q < \infty$. An exponent below one gives a quasi-norm. We apply the theorem to the additive criteria obtained by raising each norm or quasi-norm to its defining power. All sequence spaces in this subsection are real.

Let I be finite or $I = \mathbb{N}$, let $d_i \in \mathbb{R} \setminus \{0\}$, and fix $f \in \ell^p(I)$, where $0 < p, q < \infty$. Write $(Dg)_i = d_i g_i$ and set

$$C(g) = \sum_{i \in I} |g_i|^q = \|g\|_q^q, \qquad E(g) = \sum_{i \in I} |f_i - d_i g_i|^p = \|f - Dg\|_p^p, \tag{A.1}$$

$$F_{p,q}(B) = \inf_{\substack{g \in \ell^q(I) \\ C(g) \le B}} E(g), \qquad \mathcal{V}_{p,q}(\sigma) = \inf_{g \in \ell^q(I)} \{\|f - Dg\|_p + \sigma \|g\|_q\}. \tag{A.2}$$

The residual may be infinite for some competitors. Both infima are finite, however, because $g = 0$ is admissible. No boundedness of D from $\ell^q(I)$ to $\ell^p(I)$ is required.

Proposition A.1 (A model for diagonal approximation)**.** *For the data above, there are a finite or countable set J, positive weights A_j, W_j with $\sum_j W_j = \|f\|_p^p/2$, and maps*

$$P \colon \ell^q(I) \longrightarrow \Xi_A, \qquad L \colon \Xi_A \longrightarrow \ell^q(I), \qquad \Xi_A = \{\xi \in \{0,1\}^J : \sum_j A_j \xi_j < \infty\},$$

such that, with $C_d(\xi) = \sum_j A_j \xi_j$ and $E_d(\xi) = \sum_j W_j (1 - \xi_j)$,

$$C_d(Pg) \le \|g\|_q^q, \qquad E_d(Pg) \le \|f - Dg\|_p^p, \tag{A.3}$$

$$\|L\xi\|_q^q \le 2 C_d(\xi), \qquad \|f - DL\xi\|_p^p \le 2 E_d(\xi). \tag{A.4}$$

The minima defining $F_{p,q}(B)$ and $D(B) = \min_{C_d(\xi) \le B} E_d(\xi)$ are attained, and

$$D(B) \le F_{p,q}(B) \le 2 D(B/2), \qquad B \ge 0. \tag{A.5}$$

In particular, for a prescribed bound $R \geq 0$ on $\|g\|_q$,

$$D(R^q)^{1/p} \leq \inf_{\|g\|_q \leq R} \|f - Dg\|_p \leq 2^{1/p} D(R^q/2)^{1/p}. \tag{A.6}$$

Proof. Take the component choices $u \in \mathbb{R}$ and criteria $C_i(u) = |u|^q$, $E_i(u) = |f_i - d_i u|^p$. The optimal error functions are

$$F_i(B) = (|f_i| - |d_i| B^{1/q})_+^p. \tag{A.7}$$

Indeed, $|u| \leq B^{1/q}$ gives $|f_i - d_i u| \geq (|f_i| - |d_i| B^{1/q})_+$ by the reverse triangle inequality. Equality follows by taking $d_i u$ on the segment joining 0 to f_i, with magnitude $\min\{|f_i|, |d_i| B^{1/q}\}$. This also covers $f_i = 0$.

To prove that the upward component diagram is closed, let (s_n, t_n) in that diagram converge to (s, t) and choose u_n with $|u_n|^q \leq s_n$ and $|f_i - d_i u_n|^p \leq t_n$. The sequence u_n is bounded. Pass to a convergent subsequence and use continuity of both criteria to obtain an admissible choice with bounds (s, t). The unique choice with $C_i(u) = 0$ is $u = 0$, its error is $e_i = |f_i|^p$, and zero error is attained at $u = f_i/d_i$. Thus all component assumptions of Theorem 7.4 hold, with $\sum_i e_i = \|f\|_p^p < \infty$. Apply that theorem with $a = b = 2$. The domain on which C is finite is precisely $\ell^q(I)$, and its total criteria are (A.1). The theorem gives the weights and maps, proves attainment, and establishes (A.5). Substituting $B = R^q$ and taking pth roots gives (A.6). When $f = 0$, take the empty model and let L be the zero map. □

Remark A.2 (Zero diagonal entries). If some d_i vanish, set $I_0 = \{i : d_i = 0\}$ and $\beta = \sum_{i \in I_0} |f_i|^p$. Setting $g_i = 0$ on I_0 does not increase $\|g\|_q^q$ and leaves the error unchanged. Apply the construction to $I \setminus I_0$ and add the fixed error β. The optimal error for the full powered criteria then satisfies

$$\beta + D(B) \leq F_{p,q}(B) \leq \beta + 2D(B/2).$$

For the original objective, the full residual is $(\beta + E(g))^{1/p}$, where E now denotes the error on $I \setminus I_0$; the corresponding model residual is $(\beta + E_d(\xi))^{1/p}$. The fixed residual error is added before taking the root.

Corollary A.3 (The original penalized problem)**.** *For $\sigma > 0$, define*

$$R_d(\sigma) = \inf_{\xi \in \Xi_A} \{E_d(\xi)^{1/p} + \sigma C_d(\xi)^{1/q}\}, \qquad \gamma_{p,q} = \max\{2^{1/p}, 2^{1/q}\}. \tag{A.8}$$

Then the same model gives, for every $\sigma > 0$,

$$R_d(\sigma) \leq \mathcal{V}_{p,q}(\sigma) \leq 2^{1/p} R_d(2^{1/q - 1/p}\sigma) \leq \gamma_{p,q} R_d(\sigma). \tag{A.9}$$

Proof. Apply Lemma 2.1 to the coordinatewise nondecreasing objective

$$\Psi_\sigma(c, e) = e^{1/p} + \sigma c^{1/q}.$$

The separate estimates give the constants explicitly. Taking roots in (A.3) shows that the model value at Pg is at most $\|f - Dg\|_p + \sigma \|g\|_q$. Taking the infimum over g gives the first inequality. Conversely, (A.4) gives

$$\|f - DL\xi\|_p + \sigma \|L\xi\|_q \leq 2^{1/p}(E_d(\xi)^{1/p} + 2^{1/q - 1/p}\sigma C_d(\xi)^{1/q}).$$

Taking the infimum over ξ proves the second inequality. For each ξ, its right-hand side is at most $\gamma_{p,q}(E_d(\xi)^{1/p} + \sigma C_d(\xi)^{1/q})$. Taking infima once more proves the last inequality. All values are finite because the zero binary vector has finite error and satisfies $C_d(0) = 0$. □

The roots in (A.8) correspond to the two norms or quasi-norms in the original problem. The additive objective $E_d + \sigma C_d$ corresponds to penalizing the powered criteria in (A.1).

Corollary A.4 (Bounds from a finite diagonal model)**.** *Let $Q \subseteq J$ be finite, let $\omega_Q = \sum_{j \notin Q} W_j$, and let D_Q be the finite model in (7.19). Then, for every $B \geq 0$,*

$$\max\{0, D_Q(B) - \omega_Q\} \leq F_{p,q}(B) \leq 2D_Q(B/2). \tag{A.10}$$

An optimal binary vector supported in Q with $C_d(\xi) \leq B/2$ gives a sequence g_Q such that

$$\|g_Q\|_q \leq B^{1/q}, \qquad \|f - Dg_Q\|_p \leq \left(2D_Q(B/2)\right)^{1/p}. \tag{A.11}$$

For the original penalty, set

$$R_Q(\sigma) = \min_{S \subseteq Q} \left\{ \left(\omega_Q + \sum_{j \in Q \setminus S} W_j \right)^{1/p} + \sigma \left(\sum_{j \in S} A_j \right)^{1/q} \right\}. \tag{A.12}$$

If $M_d = \sum_j W_j > 0$, define

$$\eta_p(\omega) = \begin{cases} \omega^{1/p}, & p \geq 1, \\ p^{-1} M_d^{1/p-1} \omega, & 0 < p < 1. \end{cases}$$

For $M_d = 0$ put $\eta_p(0) = 0$. Uniformly for $\sigma > 0$,

$$0 \leq R_Q(\sigma) - R_d(\sigma) \leq \eta_p(\omega_Q), \tag{A.13}$$

$$\max\{0, R_Q(\sigma) - \eta_p(\omega_Q)\} \leq \mathcal{V}_{p,q}(\sigma) \leq 2^{1/p} R_Q(2^{1/q-1/p}\sigma). \tag{A.14}$$

Proof. Proposition 7.5, applied to Proposition A.1, gives (A.10). Extend an optimal binary vector by zero outside Q and apply the reconstruction map L. Its full model error is $D_Q(B/2)$, which includes all omitted weights. The two bounds in (A.4) give (A.11) after taking roots.

For (A.13), note that binary vectors supported in Q form a subset of Ξ_A, so $R_Q \geq R_d$. Given any $\xi \in \Xi_A$, set $\xi_j^Q = \xi_j$ on Q and $\xi_j^Q = 0$ otherwise. Then

$$C_d(\xi^Q) \leq C_d(\xi), \qquad 0 \leq E_d(\xi^Q) - E_d(\xi) = \sum_{j \notin Q} W_j \xi_j \leq \omega_Q.$$

Both errors lie in $[0, M_d]$. If $p \geq 1$, subadditivity of $t^{1/p}$ bounds the increase after taking roots by $\omega_Q^{1/p}$. If $0 < p < 1$, the mean value theorem and $(t^{1/p})' \leq p^{-1} M_d^{1/p-1}$ on that interval give the stated bound. The term involving C_d cannot increase. Consequently the value at ξ^Q is at most the value at ξ plus $\eta_p(\omega_Q)$. Taking infima over ξ proves (A.13). Combining it with (A.9) gives (A.14). □

Example A.5 (A diagonal problem without a square-summable exact inverse)**.** Let $I = \mathbb{N}$, $p = q = 2$ and $d_i = f_i = 2^{-i}$. Then $f \in \ell^2$, but $Dg = f$ forces $g_i = 1$ for every i, which is not in ℓ^2. The component criteria and optimal error functions are

$$C_i(u) = u^2, \qquad E_i(u) = 4^{-i}(1-u)^2, \qquad F_i(B) = 4^{-i}(1 - \sqrt{B})_+^2.$$

The geometric construction can be computed explicitly. More generally, for one nonzero datum with $p = q = 2$, write $T_i = (f_i/d_i)^2$, $e_i = f_i^2$, and

$$c = (1 - 1/\sqrt{2})^2 = \frac{3}{2} - \sqrt{2}.$$

Then $F_i(B) = e_i(1 - \sqrt{B/T_i})_+^2$. With $a = b = 2$, the construction in Lemma 7.3 chooses the nonterminal points

$$(t_1, u_1) = (cT_i, e_i/2), \qquad (t_2, u_2) = (T_i/4, e_i/4), \qquad (t_3, u_3) = (T_i/2, ce_i),$$

followed by $(T_i, 0)$. Indeed, the initial threshold satisfies $F_i(cT_i) = e_i/2$. Since $F_i(2cT_i) = 4ce_i > e_i/4$, the next point occurs at the threshold $T_i/4$. Doubling $T_i/4$ gives $F_i(T_i/2) = ce_i < e_i/8$, which determines the third point. Doubling once more gives T_i, where the error vanishes. This final point adds no further model coordinate, so the row has exactly three weights:

$$(A_{i,1}, A_{i,2}, A_{i,3}) = T_i \left(c, \frac{1}{4} - c, \frac{1}{4}\right), \qquad (W_{i,1}, W_{i,2}, W_{i,3}) = e_i \left(\frac{1}{4}, \frac{1}{4} - c, c\right). \tag{A.15}$$

The weights $A_{i,k}$ increase and the weights $W_{i,k}$ decrease. Any selected subset can therefore be replaced by the initial segment of the same cardinality without increasing either criterion. On these initial segments, we may define the reconstruction map by

Number selected k	$C_d(k)/T_i$	$E_d(k)/e_i$	$L_i(k)/(f_i/d_i)$
0	0	1/2	0
1	c	1/4	$\sqrt{2} - 1$
2	1/4	c	$1/\sqrt{2}$
3	1/2	0	1

For $k \geq 1$ these amplitudes minimize the original error under $C_i(u) \leq 2C_d(k)$. They satisfy $C_i(u) = 2C_d(k)$; their errors are respectively $4ce_i$, ce_i, and zero, which are bounded by $2E_d(k)$. The empty selection gives zero, with error $e_i = 2E_d(0)$.

In the present example $T_i = 1$ and $e_i = 4^{-i}$, so (A.15) and the table specify the model and its reconstruction map. In particular, the set Q_N consisting of the three coordinates in each of the first N rows has $3N$ binary variables, and its exact omitted weight is

$$\omega_{Q_N} = \frac{1}{2} \sum_{i>N} 4^{-i} = \frac{4^{-N}}{6}. \tag{A.16}$$

For each finite B, the minimum $F_{2,2}(B)$ is attained by Proposition A.1. The minimum is strictly positive, since a value of zero would give an exact inverse in the prescribed space. For the elementary truncation $g_i^{(N)} = 1$ when $i \leq N$ and zero otherwise, direct summation gives

$$\|g^{(N)}\|_2^2 = N, \qquad \|f - Dg^{(N)}\|_2^2 = \sum_{i>N} 4^{-i} = \frac{4^{-N}}{3}. \tag{A.17}$$

Thus $F_{2,2}(B) \to 0$ as $B \to \infty$, although it never vanishes at a finite B. If $F_{2,2}^{[N]}$ denotes the original minimum with $g_i = 0$ for $i > N$, Corollary 7.6 gives

$$0 \leq F_{2,2}^{[N]}(B) - F_{2,2}(B) \leq \frac{4^{-N}}{3}, \qquad B \geq 0. \tag{A.18}$$

For the upper bound, one can also replace the tail of a minimizer g by zero. Its squared ℓ^2 norm does not increase, and the error increases by at most $\sum_{i>N} 4^{-i}$. The lower bound follows by restricting the admissible set. Here $M_d = \|f\|_2^2/2 = 1/6$ and $\eta_2(\omega_{Q_N}) = 2^{-N}/\sqrt{6}$. In addition to the component truncation (A.18), Corollary A.4 gives finite binary models and admissible sequences satisfying the bounds (A.10)–(A.14).

For fixed data in Proposition A.1, a single set of weights and reconstruction maps applies for every bound B and every penalty parameter. They produce approximating sequences with the prescribed bound on their norm or quasi-norm, together with finite approximations to the original penalized minimum. The argument allows countably many coordinates and the ranges $p < 1$ or $q < 1$. As Q increases to J, both ω_Q and $\eta_p(\omega_Q)$ tend to zero, uniformly in B and the penalty parameter. The truncation error therefore vanishes. The factors $(2, 2)$ in the powered criteria, and the corresponding factors after taking roots, remain fixed.

A.2 Countably many smoothing problems

Smoothing splines are among the motivating examples in [10, Section 1]. The theorem for infinitely many choices applies to a countable family of independent smoothing problems with a common bound on the sum of their squared derivative seminorms.

Proposition A.6 (Countable smoothing problems)**.** *For each $i \in \mathbb{N}$, let $a_i < b_i$, let $r_i \geq 1$ be an integer, and let $a_i \leq t_{i1} < \cdots < t_{in_i} \leq b_i$, where $n_i \geq r_i$. Given data $y_{ij} \in \mathbb{R}$ and weights $w_{ij} > 0$, define, for $f \in H^{r_i}(a_i, b_i)$,*

$$C_i(f) = \int_{a_i}^{b_i} |f^{(r_i)}(t)|^2 \, dt, \qquad E_i(f) = \sum_{j=1}^{n_i} w_{ij} |f(t_{ij}) - y_{ij}|^2, \tag{A.19}$$

$$e_i = \min_{p \in \mathcal{P}_{r_i - 1}} \sum_{j=1}^{n_i} w_{ij} |p(t_{ij}) - y_{ij}|^2, \qquad \sum_{i=1}^{\infty} e_i < \infty. \tag{A.20}$$

Here $\mathcal{P}_{r_i - 1}$ denotes the polynomials of degree at most $r_i - 1$. Define

$$F(B) = \inf \left\{ \sum_i E_i(f_i) : f_i \in H^{r_i}(a_i, b_i), \ \sum_i C_i(f_i) \leq B \right\}. \tag{A.21}$$

The minimum in (A.21) is attained for every $B \geq 0$. Theorem 7.4 supplies a countable weighted model, with index set Ω and weights $(A_\nu, W_\nu)_{\nu \in \Omega}$, such that

$$\sum_{\nu \in \Omega} W_\nu = \frac{1}{2} \sum_i e_i, \qquad D(B) \leq F(B) \leq 2D(B/2). \tag{A.22}$$

For a finite set $Q \subset \Omega$, write

$$\omega_Q = \sum_{\nu \notin Q} W_\nu, \qquad D_Q(b) = \omega_Q + \min_{\substack{S \subset Q \\ \sum_{\nu \in S} A_\nu \leq b}} \sum_{\nu \in Q \setminus S} W_\nu. \tag{A.23}$$

For every $B \geq 0$, there are functions $f_{i,B,Q} \in H^{r_i}(a_i, b_i)$ with

$$\sum_i C_i(f_{i,B,Q}) \leq B \tag{A.24}$$

and the error bounds

$$\begin{aligned} \max\{0, D_Q(B) - \omega_Q\} \leq F(B) &\leq \sum_i E_i(f_{i,B,Q}) \\ &\leq 2D_Q(B/2) \leq 2F(B/2) + 2\omega_Q. \end{aligned} \tag{A.25}$$

Only finitely many of these functions differ from the fixed polynomials with $C_i = 0$. For fixed Q, all nonpolynomial choices belong to a finite collection of component representatives.

Proof. Fix a component and omit its index. Point evaluation is continuous on $H^r(a,b)$ because $r \geq 1$. We first prove the estimate

$$\|f\|_{H^r(a,b)} \leq c \left(\|f^{(r)}\|_{L^2(a,b)} + \sum_{j=1}^{r} |f(t_j)| \right). \tag{A.26}$$

The integral form of Taylor's formula for Sobolev functions gives $f = p + v$, where $p \in \mathcal{P}_{r-1}$ and

$$v(t) = \frac{1}{(r-1)!} \int_a^t (t-s)^{r-1} f^{(r)}(s)\, ds.$$

Differentiating this formula up to order r, and applying Cauchy–Schwarz to each of its integral terms, gives $\|v\|_{H^r} + \max_j |v(t_j)| \leq c_1 \|f^{(r)}\|_2$. The values of p at the first r distinct nodes determine its coefficients, since the corresponding Vandermonde matrix is invertible. Consequently $\|p\|_{H^r} \leq c_2 \sum_{j=1}^{r} |p(t_j)|$, which proves (A.26).

Suppose now that (b_k, d_k) belongs to the upward attainable diagram and tends to a finite pair (b,d). Choose f_k with $C(f_k) \leq b_k$ and $E(f_k) \leq d_k$. Positivity of the weights gives

$$|f_k(t_j)| \leq |y_j| + w_j^{-1/2} d_k^{1/2}.$$

Estimate (A.26) therefore bounds (f_k) in H^r. After passage to a subsequence, $f_k \rightharpoonup f$ in H^r. Weak lower semicontinuity applies to the derivative norm, and the point values converge because point evaluation is a bounded linear functional. Thus $C(f) \leq b$ and $E(f) \leq d$, proving closedness of the upward diagram.

The equality $C(f) = 0$ holds precisely for $f \in \mathcal{P}_{r-1}$. The sample map is injective on this finite-dimensional space, since there are at least r distinct nodes. Its weighted least-squares problem therefore has a minimizer p, with error e as defined in (A.20). On the other hand, a polynomial interpolating all n data has error zero and finite derivative seminorm. Hence $\inf E = 0$. Components with $e_i = 0$ may be fixed at their minimizing polynomials in $\mathcal{P}_{r_i-1}$ and omitted. The hypotheses of Theorem 7.4 are now satisfied; its summability assumption is exactly (A.20). We obtain (A.22).

Choose a minimizing binary vector for $D_Q(B/2)$ and apply the reconstruction map of Theorem 7.4. Its separate bounds for the two criteria give (A.24) and $\sum_i E_i(f_{i,B,Q}) \leq 2D_Q(B/2)$. Proposition 7.5 gives $0 \leq D_Q(b) - D(b) \leq \omega_Q$. Combining this with $D(b) \leq F(b)$ proves the remaining inequalities in (A.25). The reconstruction chooses among the retained component representatives. In every component outside the support of this vector it uses the fixed polynomial with $C_i = 0$. Since the support meets only finitely many components, this proves the final assertion. □

This construction approximates a family of data sets under a bound on the sum of their squared derivative seminorms. The summability assumption concerns only the errors of the best fits by polynomials of degree at most $r_i - 1$. These polynomial parts may be large because their r_ith derivatives vanish. A finite model improves the fit in finitely many components. The reconstructed functions satisfy the prescribed bound on the sum of squared seminorms and the error estimate (A.25).

Example A.7 (Explicit fits with a bound on the derivatives)**.** For each $i \in \mathbb{N}$, take the real Sobolev space $H^1(0,1)$, the two nodes $0, 1$, data $y_{i1} = m_i - 1/2$, $y_{i2} = m_i + 1/2$, and equal weights $w_{i1} = w_{i2} = 2\,4^{-i}$, where the means $m_i \in \mathbb{R}$ are arbitrary. Thus

$$C_i(f) = \int_0^1 |f'(t)|^2\, dt, \qquad E_i(f) = 2\,4^{-i} \left(|f(0) - m_i + 1/2|^2 + |f(1) - m_i - 1/2|^2 \right).$$

The constant fit $f = m_i$ satisfies $C_i(f) = 0$ and has error $e_i = 4^{-i}$. In particular, $\sum_i e_i = 1/3$, with no summability requirement on the means.

The component optimal error function can be computed explicitly. For any $f \in H^1(0,1)$, put $s = f(1) - f(0)$ and $h = (f(0) + f(1))/2 - m_i$. Cauchy–Schwarz and the identity $s = \int_0^1 f'(t)\,dt$ give $C_i(f) \geq s^2$, while

$$E_i(f) = 4^{-i}(4h^2 + (1-s)^2).$$

For $C_i(f) \leq b$, the least possible value of $(1-s)^2$ is $(1-\sqrt{b})_+^2$. It is attained with $h = 0$ by the affine function $f(t) = m_i + \min\{1, \sqrt{b}\}(t - 1/2)$. Hence

$$F_i(b) = 4^{-i}(1-\sqrt{b})_+^2, \qquad b \geq 0. \tag{A.27}$$

Every exact fit satisfies $C_i(f) \geq 1$, so fitting all components exactly would make the sum of squared derivative seminorms infinite.

With $c = 3/2 - \sqrt{2}$ as in Example A.5, the three weights in each row are therefore

$$(A_{i,1}, A_{i,2}, A_{i,3}) = \left(c, \frac{1}{4} - c, \frac{1}{4}\right), \qquad (W_{i,1}, W_{i,2}, W_{i,3}) = 4^{-i}\left(\frac{1}{4}, \frac{1}{4} - c, c\right). \tag{A.28}$$

All the required representatives are the four affine functions

$$f_{i,k}(t) = m_i + \rho_k(t - 1/2), \qquad (\rho_0, \rho_1, \rho_2, \rho_3) = (0, \sqrt{2} - 1, 1/\sqrt{2}, 1). \tag{A.29}$$

For an initial segment of length k, the two criteria are

$$(C_i(f_{i,k}))_{k=0}^3 = \left(0, 2c, \frac{1}{2}, 1\right), \qquad (E_i(f_{i,k}))_{k=0}^3 = 4^{-i}(1, 4c, c, 0).$$

Their values of C_i are twice the corresponding partial sums of the weights $A_{i,j}$, and their errors are at most twice the sum of the remaining row weights.

Let Q_N contain the three coordinates in each of the first N rows. Its omitted weight is $\omega_{Q_N} = 4^{-N}/6$. Replacing a row selection by an initial segment does not increase either model criterion, so the finite minimum can be written as

$$D_{Q_N}(b) = \frac{4^{-N}}{6} + \min_{\substack{k_i \in \{0,1,2,3\},\ 1 \leq i \leq N \\ \sum_{i=1}^N b_{k_i} \leq b}} \sum_{i=1}^N 4^{-i}\varepsilon_{k_i}, \tag{A.30}$$

where

$$(b_0, b_1, b_2, b_3) = \left(0, c, \frac{1}{4}, \frac{1}{2}\right), \qquad (\varepsilon_0, \varepsilon_1, \varepsilon_2, \varepsilon_3) = \left(\frac{1}{2}, \frac{1}{4}, c, 0\right).$$

Choose a minimizing tuple in (A.30) at $b = B/2$. Use its representative f_{i,k_i} for $i \leq N$ and the constant m_i for $i > N$, and denote this family by $(f_{i,B,N})_i$. The sum of its squared derivative seminorms is at most B, and

$$\begin{aligned} \max\left\{0, D_{Q_N}(B) - \frac{4^{-N}}{6}\right\} &\leq F(B) \leq \sum_i E_i(f_{i,B,N}) \\ &\leq 2D_{Q_N}(B/2) \leq 2F(B/2) + \frac{4^{-N}}{3}. \end{aligned} \tag{A.31}$$

Here F is the optimal error function for the full countable problem in (A.21). The same omitted weight bounds $D_{Q_N}(b) - D(b)$ for every $b \geq 0$. There are at most 4^N such families of fits, all determined before B is specified. Varying B selects among these families.

For instance, with $N = 2$ and $B = 1$, the minimizing initial segments have lengths $k_1 = k_2 = 2$. Every other feasible pair except $(3, 0)$ and $(0, 3)$ has both lengths at most two, so its remaining error is at least that of $(2, 2)$. The pair $(2, 2)$ also has smaller remaining error than $(3, 0)$ and $(0, 3)$, since $5c/16 < 1/32$. The selected functions are $f_{i,1,2}(t) = m_i + (t - 1/2)/\sqrt{2}$ for $i = 1, 2$, and m_i for $i > 2$. They satisfy

$$\sum_i C_i(f_{i,1,2}) = 1, \qquad \sum_i E_i(f_{i,1,2}) = \frac{5c}{16} + \frac{1}{48}.$$

The finite model therefore determines fitted functions under the bound on the sum of squared derivative seminorms. The error bound (A.31) holds uniformly in B.

A.3 Bias and noise amplification in linear recovery

The recovery estimate in [10, equation (1.12)] bounds the error by the sum of deterministic approximation error and amplified observational noise. We apply the theorem for infinitely many choices to the recovery operator. The model approximates the infimum of this upper bound. Its operator representatives are fixed, and the noise level determines which representative is selected.

Proposition A.8 (A model for linear recovery bounds)**.** *Let X, Y be Banach spaces, let Z be a reflexive Banach space, let $K \subset X$ be nonempty and bounded, and let $U : X \to Z$ and $I : X \to Y$ be bounded linear operators. For $R \in \mathcal{L}(Y, Z)$, define*

$$\begin{aligned} \beta(R) &= \sup_{x \in K} \|Ux - RIx\|_Z, \qquad & \mu &= \inf_{R \in \mathcal{L}(Y,Z)} \beta(R), \\ V(\delta) &= \inf_{R \in \mathcal{L}(Y,Z)} \{\beta(R) + \delta\|R\|\}, \qquad & \delta &> 0. \end{aligned} \tag{A.32}$$

There is a countable weighted model satisfying

$$\sum_{\nu \in \Omega} W_\nu = \frac{1}{2}\left(\sup_{x \in K} \|Ux\|_Z - \mu\right), \qquad M(\delta) = \sum_{\nu \in \Omega} \min\{W_\nu, \delta A_\nu\}, \tag{A.33}$$
$$\mu + M(\delta) \leq V(\delta) \leq \mu + 2M(\delta).$$

For finite $Q \subset \Omega$, put

$$\omega_Q = \sum_{\nu \notin Q} W_\nu, \qquad M_Q(\delta) = \omega_Q + \sum_{\nu \in Q} \min\{W_\nu, \delta A_\nu\}. \tag{A.34}$$

Then $0 \leq M_Q(\delta) - M(\delta) \leq \omega_Q$, and there is a bounded linear recovery operator $R_{Q,\delta} : Y \to Z$ such that

$$\sup_{\substack{x \in K \\ \|w\|_Y \leq \delta}} \|Ux - R_{Q,\delta}(Ix + w)\|_Z \leq \beta(R_{Q,\delta}) + \delta\|R_{Q,\delta}\| \tag{A.35}$$

$$\leq \mu + 2M_Q(\delta) \leq 2V(\delta) - \mu + 2\omega_Q. \tag{A.36}$$

Thus V is the infimum of the upper bound given by bias plus noise amplification. Estimate (A.36) bounds the recovery error of the constructed operators.

Proof. The boundedness assumptions make $\beta(R)$ finite for every bounded linear R, and $0 \le \mu \le \beta(0) < \infty$. Use the single-component criteria

$$C(R) = \|R\|, \qquad E(R) = \beta(R) - \mu.$$

Under $C(R) = 0$, the error is $e_0 = \beta(0) - \mu$, attained at $R = 0$, and $\inf_R E(R) = 0$ by the definition of μ.

We prove closedness by using compactness for pointwise weak convergence. For $L < \infty$, embed the ball $\{R \in \mathcal{L}(Y, Z) : \|R\| \le L\}$ in

$$\prod_{y \in Y} \{z \in Z : \|z\|_Z \le L\|y\|_Y\},$$

giving each factor its weak topology. Reflexivity and Tychonoff's theorem make this product compact. The operator ball is a closed subset: a pointwise weak limit preserves all linear identities, and its coordinate bounds imply $\|Ry\| \le L\|y\|$. It is therefore compact for pointwise weak convergence.

Let (c_k, e_k) be attainable upper bounds tending to (c, e), and choose R_k with $\|R_k\| \le c_k$ and $\beta(R_k) \le \mu + e_k$. All these operators lie in one such ball. A subnet converges pointwise weakly to a bounded linear operator R. Weak lower semicontinuity gives, for each $y \in Y$ and $x \in K$,

$$\|Ry\| \le c\|y\|, \qquad \|Ux - RIx\| \le \mu + e.$$

Taking the supremum over x proves $C(R) \le c$ and $E(R) \le e$. The upward diagram is closed, so Theorem 7.4 applies. If $e_0 = 0$, take an empty model and $R_{Q,\delta} = 0$ throughout.

For $e_0 > 0$, the minimum of the model penalty is $M(\delta)$: select precisely the coordinates with $\delta A_\nu < W_\nu$. The sum of the selected weights A_ν is finite, since

$$\sum_{\delta A_\nu < W_\nu} A_\nu \le \delta^{-1} \sum_{\nu} W_\nu < \infty.$$

For each recovery operator, the forward map gives a model selection whose two criteria do not exceed $\|R\|$ and $\beta(R) - \mu$, respectively. Taking infima gives $M(\delta) \le V(\delta) - \mu$. Applying the reconstruction map to a minimizing selection gives $V(\delta) - \mu \le 2M(\delta)$ and proves (A.33).

For finite Q, take the selection $S_{Q,\delta} = \{\nu \in Q : \delta A_\nu < W_\nu\}$ and let $R_{Q,\delta}$ be its image under the reconstruction map. The omitted coordinates contribute their full error weight ω_Q. The separate reconstruction bounds therefore give $\beta(R_{Q,\delta}) + \delta\|R_{Q,\delta}\| \le \mu + 2M_Q(\delta)$. Also

$$M_Q(\delta) - M(\delta) = \sum_{\nu \notin Q} (W_\nu - \min\{W_\nu, \delta A_\nu\}) \in [0, \omega_Q].$$

Together with $M(\delta) \le V(\delta) - \mu$, this proves (A.36). Finally, linearity and the triangle inequality yield

$$\|Ux - R(Ix + w)\| \le \|Ux - RIx\| + \|R\|\|w\| \le \beta(R) + \delta\|R\|,$$

which proves (A.35). □

The same model applies throughout the range $\delta > 0$. For fixed Q, it gives a finite family of bounded linear recovery operators; the noise level determines which operator to use. The omitted weight ω_Q bounds the error introduced by truncating the model, uniformly in the noise level. In all three applications, the maps from the model produce admissible approximations and separate estimates for the two criteria. Exhausting the model by finite subsets removes the truncation error. The comparison factors remain those of the variational fundamental lemma.

References

[1] C. Bazgan, S. Ruzika, C. Thielen, and D. Vanderpooten, *The power of the weighted sum scalarization for approximating multiobjective optimization problems*, Theory of Computing Systems **66** (2022), 395–415. doi:10.1007/s00224-021-10066-5.

[2] A. Cohen, R. DeVore, and E. Tadmor, *Constructions of bounded solutions of* $\operatorname{div} \mathbf{u} = f$ *in critical spaces*, in *Multiscale, Nonlinear and Adaptive Approximation II*, Springer, 2024, 177–200. doi:10.1007/978-3-031-75802-7_9.

[3] M. Cwikel, B. Jawerth, and M. Milman, *On the fundamental lemma of interpolation theory*, Journal of Approximation Theory **60** (1990), 70–82. doi:10.1016/0021-9045(90)90074-Z.

[4] O. Domínguez and M. Milman, *Quantitative* Γ*-convergence: a variational extrapolation framework*, preprint, 2026. arXiv:2610.06645.

[5] A. Federgruen and H. Groenevelt, *The greedy procedure for resource allocation problems: necessary and sufficient conditions for optimality*, Operations Research **34** (1986), no. 6, 909–918. doi:10.1287/opre.34.6.909.

[6] E. Gagliardo, *Una struttura unitaria in diverse famiglie di spazi funzionali. I*, Ricerche di Matematica **10** (1961), 244–281.

[7] E. Gagliardo, *A common structure in various families of functional spaces. II. Quasilinear interpolation spaces*, Ricerche di Matematica **12** (1963), 87–107.

[8] P. Halffmann, S. Ruzika, C. Thielen, and D. Willems, *A general approximation method for bicriteria minimization problems*, Theoretical Computer Science **695** (2017), 1–15. doi:10.1016/j.tcs.2017.07.003. Author preprint, arXiv:1701.02989.

[9] B. Jawerth and M. Milman, *Paseky Lectures on Interpolation, Image Processing and Optimization*, unpublished expanded manuscript based on the Paseky 2007 lectures, May–June 2007, 89 pp. Only the introduction appeared in the Charles University booklet; the full lecture manuscript was not published.

[10] C. A. Micchelli and A. Pinkus, *Variational problems arising from balancing several error criteria*, Rendiconti di Matematica e delle sue Applicazioni (7) **14** (1994), 37–86. Author-hosted text.

[11] M. Milman, *Interpolation theory and the calculus of variations: stable representations under Mosco convergence*, manuscript, 2026.

[12] K. Modin, A. Nachman, and L. Rondi, *A multiscale theory for image registration and nonlinear inverse problems*, Advances in Mathematics **346** (2019), 1009–1066. doi:10.1016/j.aim.2019.02.014. Corrected author version: arXiv:1803.01790v2, 2020.

[13] C. H. Papadimitriou and M. Yannakakis, *On the approximability of trade-offs and optimal access of Web sources*, Proceedings of the 41st Annual Symposium on Foundations of Computer Science, IEEE Computer Society, 2000, 86–92. doi:10.1109/SFCS.2000.892068.

[14] J. Peetre and G. Sparr, *Interpolation of normed Abelian groups*, Annali di Matematica Pura ed Applicata (4) **92** (1972), 217–262. doi:10.1007/BF02417949.

[15] G. Pisier, *The* K_t*-functional for the interpolation couple* $L^1(A_0), L^\infty(A_1)$, Journal of Approximation Theory **73** (1993), 106–117. doi:10.1006/jath.1993.1032.

[16] E. Tadmor, *Hierarchical construction of bounded solutions in critical regularity spaces*, Communications on Pure and Applied Mathematics **69** (2016), 1087–1109. doi:10.1002/cpa.21575.

[17] E. Tadmor, S. Nezzar, and L. Vese, *A multiscale image representation using hierarchical* (BV, L^2) *decompositions*, Multiscale Modeling & Simulation **2** (2004), 554–579. doi:10.1137/030600448.

Mario Milman (corresponding author)
Instituto Argentino de Matemática
Email address: mario.milman@icloud.com
URL: `https://sites.google.com/site/mariomilman`