A Categorical Representation of Population-Based Metaheuristics

Mariano Frutos
ORCID: 0000-0003-2585-4195
Department of Engineering, Universidad Nacional del Sur, Bahía Blanca, Argentina
IIESS-CONICET, Bahía Blanca, Argentina
E-mail: marianouns@gmail.com

Fernando Tohmé
ORCID: 0000-0003-2988-4519
Department of Economics, Universidad Nacional del Sur, Bahía Blanca, Argentina
INMABB-CONICET, Bahía Blanca, Argentina
E-mail: ftohme@criba.edu.ar

Abstract. Population-based metaheuristics are typically designed as monolithic procedures, making it difficult to attribute performance to specific algorithmic components. This paper develops a compositional categorical framework in which primitive modules are modeled as finite stochastic maps in **FinStoch** and complete algorithms are assembled through wiring diagrams in a symmetric monoidal category. A lax monoidal behavior functor $\mathsf{Bhv} : \mathsf{W} \to$ **FinStoch** assigns a composite stochastic kernel to each wiring diagram, so that pipeline behavior is determined by the behaviors of its constituent modules. The framework is instantiated for NSGA-II, SPEA2, and PESA-II, as well as for a two-stage MOEA + TOPSIS decision pipeline. A spectral composition theorem then bounds the spectral gap of a composed kernel in terms of its component gaps, yielding convergence-rate bounds and a finite-horizon admissibility result. A rework-production case study grounds the framework in empirical data and illustrates how local routing probabilities can be embedded into a global behavioral model. The resulting perspective supports modular comparison, diagnosis, and synthesis of stochastic optimization pipelines.

Keywords: category theory; compositional semantics; finite stochastic maps; multi-objective evolutionary algorithms; spectral gap.

Una representación categórica de metaheurísticas basadas en poblaciones

Resumen. Las metaheurísticas basadas en poblaciones suelen diseñarse como procedimientos monolíticos, lo que dificulta atribuir su desempeño a componentes algorítmicos específicos. Este artículo desarrolla un marco categórico composicional en el que los módulos primitivos se modelan como mapas estocásticos finitos en **FinStoch** y los algoritmos completos se ensamblan mediante diagramas de cableado en una categoría monoidal simétrica. Un funtor laxo monoidal de comportamiento $\mathsf{Bhv} : \mathsf{W} \to \mathbf{FinStoch}$ asigna a cada diagrama de cableado un núcleo estocástico compuesto, de modo que el comportamiento de una tubería queda determinado por los comportamientos de sus módulos constitutivos. El marco se instancia para NSGA-II, SPEA2 y PESA-II, así como para una tubería de decisión MOEA + TOPSIS en dos etapas. Luego, un teorema de composición espectral acota la brecha espectral de un núcleo compuesto en términos de las brechas de sus componentes, obteniendo cotas de convergencia y un resultado de admisibilidad de horizonte finito. Un estudio de caso sobre producción con retrabajo vincula el marco con datos empíricos.

Palabras clave: teoría de categorías; semántica composicional; mapas estocásticos finitos; algoritmos evolutivos multiobjetivo; brecha espectral.

1 Introduction

Population-based metaheuristics, including genetic algorithms, evolutionary strategies, particle-swarm methods, and multi-objective evolutionary algorithms (MOEAs), are widely deployed for hard combinatorial and continuous optimization problems (Talbi, 2009). Their success rests on a conceptual fragility: predicting, or explaining, pipeline behavior from component behaviors is generally hard. When an MOEA outperforms a competitor on a benchmark suite, it is often unclear whether the advantage arises from selection, variation, archiving, or hyperparameter tuning alone.

Theoretical frameworks for randomized search heuristics analyze individual algorithms as monolithic objects (Auger and Doerr, 2011), while practitioners assemble algorithms from reusable components, including hybridizations, memetic wrappers, and surrogate models, at a pace that outstrips compositional theory. Category theory, and specifically wiring diagrams and hypergraph categories, provides a principled answer via a lax monoidal behavior functor: a structure-preserving map assigning behaviors to wiring diagrams in a way that respects both sequential and parallel composition (Baez and Fong, 2015; Fong and Spivak, 2019; Tohmé and Rossit, 2025).

Our central contributions are: (i) primitive modules (evaluation, selection, variation, archiving) as finite stochastic maps, i.e., morphisms in $\mathbf{Kl}(\mathcal{D})$; (ii) full algorithms as wiring diagrams in a symmetric monoidal category; (iii) a lax monoidal functor $\mathsf{Bhv} : \mathsf{W} \to \mathbf{FinStoch}$; (iv) a spectral mixing-time composition theorem bounding the spectral gap of a composed kernel; and (v) finite-horizon admissibility as a corollary of spectral convergence. The framework is instantiated for NSGA-II, SPEA2, and PESA-II (Corne et al., 2001; Zitzler et al., 2001; Deb et al., 2002), and for the MOEA + TOPSIS pipeline of Miguel et al. (2024).

2 Mathematical Preliminaries

Definition 2.1 (**FinStoch**). *The category* **FinStoch** *has finite sets as objects and stochastic matrices $k : Y \times X \to [0,1]$ with $\sum_y k(y \mid x) = 1$ as morphisms $k : X \to Y$. Composition is matrix multiplication; identities are Kronecker deltas. It is isomorphic to the Kleisli category $\mathbf{Kl}(\mathcal{D})$ of the finitely supported distribution monad $\mathcal{D}$.*

FinStoch carries a symmetric monoidal structure: $X \otimes Y := X \times Y$, the tensor of morphisms is the independent product of kernels, and the unit is $I = \{*\}$. Copy $\Delta_X : X \to X \times X$ and delete $!_X : X \to \{*\}$ make it a Markov category (Fritz, 2020).

Definition 2.2 (Wiring category W). *Fix a type set Σ. Let* **TFS** *be the category of typed finite sets. Define*

$$\mathsf{W} := \mathbf{Cospan}(\mathbf{TFS}),$$

whose objects are interfaces and whose morphisms $X \to Y$ are cospans $X \xrightarrow{f} N \xleftarrow{g} Y$, called wiring diagrams. Composition is by pushout; the monoidal product is coproduct of interfaces. W is a hypergraph category (Fong, 2015).

Definition 2.3 (Behavior functor). *A behavior functor is a lax symmetric monoidal functor $\mathsf{Bhv} : \mathsf{W} \to$ **FinStoch**. For an interface X, $\mathsf{Bhv}(X)$ is a finite state space; for a wiring diagram $w : X \to Y$, $\mathsf{Bhv}(w) : \mathsf{Bhv}(X) \to \mathsf{Bhv}(Y)$ is the induced stochastic kernel.*

3 Metaheuristics as Open Stochastic Systems

Definition 3.1 (State kernel and T-step behavior). *A state kernel for a module with interface A is a stochastic matrix $K : I \times S \to \mathcal{D}(O \times S)$, where S is the finite internal state space, I the input space, and O the output space. The T-step behavior $K^{(T)} : I \times S \to \mathcal{D}(O^T \times S)$ is defined by*

$$K^{(T)}\big((o_1, \dots, o_T, s_T) \mid (i, s_0)\big) = \sum_{s_1, \dots, s_{T-1}} \prod_{t=1}^{T} K\big((o_t, s_t) \mid (i, s_{t-1})\big).$$

State kernels compose sequentially and in parallel. Given $K_1 : I_1 \times S_1 \to \mathcal{D}(O_1 \times S_1)$ and $K_2 : O_1 \times S_2 \to \mathcal{D}(O_2 \times S_2)$,

$$\begin{aligned}&(K_2 \circ K_1)\big((o_2, s_1', s_2') \mid (i, s_1, s_2)\big)\\ &\qquad = \sum_{o_1} K_2\big((o_2, s_2') \mid (o_1, s_2)\big)\, K_1\big((o_1, s_1') \mid (i, s_1)\big).\end{aligned}$$

Proposition 3.2 (Compositionality of T-step behavior). *$(K_2 \circ K_1)^{(T)}$ is determined* *by $K_1^{(T_1)}$ and $K_2^{(T_2)}$ for any $T_1 + T_2 = T$, by associativity of matrix multiplication in* **FinStoch**.

4 MOEA Modules and Their Kernels

Let X be a finite search space, $O = (f_1, f_2)$ the discretized bi-objective space, and N the population size. The primitive modules shared by NSGA-II, SPEA2, and PESA-II are:

- Eval : $\text{Pop}_N(X) \to \mathcal{D}(\text{Pop}_N(X \times O))$,
- Rank : $\text{Pop}_N(X \times O) \to \mathcal{D}(\text{Pop}_N(X \times O \times \mathbb{N}))$,
- Div : $\text{Pop}_N(X \times O \times \mathbb{N}) \to \mathcal{D}(\text{Pop}_N(X \times O \times \mathbb{N} \times \mathbb{R}))$,
- Sel, Var, Arc, Trunc (selection, variation, archive update, truncation/elitism),

all interpreted as morphisms in **FinStoch**. The three algorithms differ only in how these modules are assembled:

$$
\begin{aligned}
K_{\text{NSGA-II}} &= \text{Trunc} \circ (\text{Rank} \otimes \text{Div}) \circ \text{Eval} \circ \text{Var} \circ \text{Sel} \circ (\text{Rank} \otimes \text{Div}) \circ \text{Eval}, \\
K_{\text{SPEA2}} &= \text{Arc} \circ \text{Eval} \circ \text{Var} \circ \text{Sel}_{\text{bin}} \circ (\text{Rank}_{\text{str}} \otimes \text{Div}_{k\text{-NN}}), \\
K_{\text{PESA-II}} &= \text{Arc}_{\text{box}} \circ \text{Eval} \circ (\text{Var} \circ \text{Sel}_{\text{box}} \oplus \text{Var} \circ \text{Sel}_{\text{ind}}).
\end{aligned}
$$

TOPSIS is a degenerate kernel $K_{\text{TOPSIS}} : A \to \mathcal{D}(X \times O)$ selecting a single solution x^* from an approximate Pareto archive A (Hwang and Yoon, 1981; Miguel et al., 2024). The two-stage pipeline of Miguel et al. (2024) is the wiring diagram $w_{\text{pipe}} \in \mathsf{W}$ composing any of the three MOEAs with TOPSIS, yielding

$$
\mathsf{Bhv}(w_{\text{pipe}}) = K_{\text{TOPSIS}} \circ K_{\text{MOEA}}^{(T)}.
$$

Proposition 4.1 (Lax monoidal behavior functor for MOEAs)**.** *The assignment* $\mathsf{Bhv} : \mathsf{W} \to$ **FinStoch** *induced by the modules above is a lax symmetric monoidal functor. Laxity reflects the fact that composing independent sub-algorithms may introduce correlations invisible from each alone.*

5 Case Study: Rework-Production System

We consider a production plant with five products $P_1, \dots, P_5$, five sequential operations $O_1, \dots, O_5$, parallel machines, and a quality-control station classifying defects into five error types $E_1, \dots, E_5$. The empirical rework routing kernel for product P_p under dominant error E_e is

$$
\widehat{\kappa}_{p,e}(m \mid P_p, E_e) = \frac{h_{e,m}}{\sum_{m' \in R(p,e)} h_{e,m'}},
$$

where $h_{e,m}$ are historical rework counts and $R(p,e)$ is the set of eligible machines. Selected probabilities are shown in Table 1.

The probabilities in Table 1 are exactly the maximum-likelihood estimates of the rework routing distributions. The sequential chain $O_1 \to \cdots \to O_5 \to$ Control is

Table 1: Rework routing kernels per product and associated error type.

Product	Error	Machine	Probability
P_1	E_2	$M_{1,1}$	0.222
		$M_{3,3}$	0.556
		$M_{5,1}$	0.222
P_2	E_4	$M_{1,3}$	0.222
		$M_{2,1}$	0.778
P_3	E_2	$M_{1,2}$	0.371
		$M_{3,4}$	0.229
		$M_{5,2}$	0.400
P_4	E_3	$M_{2,2}$	0.773
		$M_{4,1}$	0.227
P_5	E_1	$M_{2,1}$	0.167
		$M_{4,3}$	0.278
		$M_{5,1}$	0.556

modeled as a wiring diagram $w_{\text{plant}} \in \mathsf{W}$, and $\mathsf{Bhv}(w_{\text{plant}})$ assigns a composite kernel tracking each item's location, rework status, and error type. Since Bhv is a functor, sub-network behaviors compose: for sub-networks $w_1 = (O_1 \to O_3)$ and $w_2 = (O_4 \to \text{Control})$, $\mathsf{Bhv}(w_2 \circ w_1) = \mathsf{Bhv}(w_2) \circ \mathsf{Bhv}(w_1)$.

6 Compositionality, Convergence, and Admissibility

6.1 Spectral mixing-time composition

Definition 6.1 (Spectral gap and mixing time)**.** *Let $P : S \to \mathcal{D}(S)$ be reversible with respect to stationary distribution π on a finite state space S. Define $\lambda_*(P) := \max\{|\lambda| : \lambda$ eigenvalue of P on $L^2_0(\pi)\}$ and the spectral gap $\gamma(P) := 1 - \lambda_*(P)$. The ε-mixing time is*

$$\tau_\varepsilon(P) := \min \left\{ t \geq 0 : \sup_{s \in S} \left\| \frac{P^t(s, \cdot)}{\pi(\cdot)} - 1 \right\|_{2,\pi} \leq \varepsilon \right\}.$$

Theorem 6.2 (Spectral mixing-time composition)**.** *Let $P_1, \dots, P_m$ be Markov kernels on the same finite state space S, each irreducible, aperiodic, and reversible with respect to a common stationary distribution π. Let $P := P_m \circ \cdots \circ P_1$. Then:*

1. *P is row-stochastic with stationary distribution π;*
2. *$\lambda_*(P) \leq \prod_{j=1}^m \lambda_*(P_j)$;*
3. *$\gamma(P) \geq 1 - \prod_{j=1}^m (1 - \gamma(P_j))$;*
4. *for every $\varepsilon \in (0, 1)$,*

$$\tau_\varepsilon(P) \leq \frac{\log(1/\varepsilon) + \frac{1}{2}\log(1/\pi_*)}{-\log\left(\prod_{j=1}^m \lambda_*(P_j)\right)}, \qquad \pi_* := \min_s \pi(s).$$

Proof. P is row-stochastic and $\pi P = \pi$ follow from closure under composition in **FinStoch** and $\pi P_j = \pi$. Since each P_j preserves $L_0^2(\pi)$ and is self-adjoint on $L_0^2(\pi)$ by reversibility, submultiplicativity of the operator norm gives

$$\|P\|_{2,\pi} \le \prod_j \|P_j\|_{2,\pi} = \prod_j \lambda_*(P_j).$$

The gap bound follows immediately, and the mixing-time bound is the standard L^2 estimate for reversible finite chains (Levin and Peres, 2017). $\square$

Corollary 6.3 (Module-wise bound for MOEA pipelines)**.** *Under the hypotheses of Theorem 6.2, any one-generation MOEA kernel obtained by composing lifted primitive modules has spectral gap bounded below by* $1 - \prod_j (1 - \gamma_j)$*, where* γ_j *is the gap of the* j*-th module.*

6.2 Observable performance and admissibility

For a finite archive $a \in A$, define metric observables: NPS $= |a|$; MID $= |a|^{-1} \sum_{x \in a} \|(f_1(x), f_2(x))\|_2$; and hypervolume HV. Each observable o_μ pushes forward the distribution $\overline{K}^{(T)}$, giving a distribution over performance values.

Definition 6.4 (Behavioral preorder)**.** *For a budget* T *and observable* o_μ*, write* $K_A \preceq_{o_\mu, I} K_B$ *if*

$$\mathbb{E}\left[o_\mu \circ \overline{K}_A^{(T)}(\cdot \mid i)\right] \le \mathbb{E}\left[o_\mu \circ \overline{K}_B^{(T)}(\cdot \mid i)\right]$$

for every instance $i \in I$*.*

Theorem 6.5 (Finite-horizon admissibility from spectral convergence)**.** *Let* K *be a lifted metaheuristic kernel satisfying the hypotheses of Theorem 6.2, with stationary distribution* π_i *and* $\lambda_i := \lambda_*(K_i)$ *per instance* $i \in I$*. Fix a target* b *and* $\delta > 0$*. If the stationary expectation satisfies* $M_{\pi_i}(o_\mu) \le b - \delta$ *for all* i*, then any horizon* T *with*

$$\sup_i \|o_\mu - M_{\pi_i}(o_\mu)\|_{2,\pi_i} \lambda_i^T \pi_i(s_{0,i})^{-1/2} \le \delta$$

guarantees $\mathbb{E}[o_\mu \circ \overline{K}^{(T)}(\cdot \mid i, s_{0,i})] \le b$ *for all* $i \in I$*, i.e.,* K *is admissible at* T*.*

Corollary 6.6 (Structural comparison of NSGA-II, SPEA2, PESA-II)**.** *If* $K_{\text{NSGA-II}} \prec_{o_{\text{HV}}, I} K_{\text{SPEA2}}$*, then the wiring diagrams of the two algorithms differ in at least one of* Rank, Div, Sel, *or* Trunc/Arc*. Hence any performance difference is attributable to at least one structurally identifiable module.*

7 Conclusion

We presented a compositional categorical semantics for population-based metaheuristics: each module is a finite stochastic map, full algorithms are wiring diagrams, and the behavior functor $\mathsf{Bhv} : \mathsf{W} \to$ **FinStoch** ensures pipeline behavior is determined compositionally. Our spectral composition theorem quantifies how module-wise convergence rates propagate to the assembled one-generation kernel, and Theorem 6.5 turns this into a finite-horizon admissibility guarantee. The rework-production case study shows how the framework connects to observable data via maximum-likelihood rework kernels. Future work includes relaxing reversibility via reversibilization techniques (Montenegro and Tetali, 2006), empirical estimation of module-wise spectral gaps, and hybrid algorithm synthesis guided by convergence bounds.

References

Auger, A., & Doerr, B. (Eds.). (2011). *Theory of randomized search heuristics*. World Scientific.

Baez, J. C., & Fong, B. (2015). A compositional framework for passive linear networks. *Theory and Applications of Categories, 33*(38), 1158-1222.

Corne, D. W., Jerram, N. R., Knowles, J. D., & Oates, M. J. (2001). PESA-II: Region-based selection in evolutionary multiobjective optimization. In *Proceedings of GECCO 2001* (pp. 283-290).

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation, 6*(2), 182-197.

Fong, B. (2015). Decorated cospans. *Theory and Applications of Categories, 30*(33), 1096-1120.

Fong, B., & Spivak, D. I. (2019). *An invitation to applied category theory*. Cambridge University Press.

Fritz, T. (2020). A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics. *Advances in Mathematics, 370*, 107239.

Hwang, C. L., & Yoon, K. (1981). *Multiple attribute decision making*. Springer.

Levin, D. A., & Peres, Y. (2017). *Markov chains and mixing times* (Vol. 107). American Mathematical Society.

Miguel, F. M., Frutos, M., Méndez, M., Tohmé, F., & González, B. (2024). Comparison of MOEAs in an optimization-decision methodology for a joint order batching and picking system. *Mathematics, 12*(8), 1246.

Montenegro, R., & Tetali, P. (2006). *Mathematical aspects of mixing times in Markov chains*. Now Publishers.

Talbi, E. G. (2009). *Metaheuristics: From design to implementation*. Wiley.

Tohmé, F., & Rossit, D. (2025). Compositional scheduling in Industry 4.0 cyber-physical systems. *Axioms, 14*(5), 332.

Zitzler, E., Laumanns, M., & Thiele, L. (2001). *SPEA2: Improving the strength Pareto evolutionary algorithm* (Technical Report 103). TIK, ETH Zürich.