ON THE GEOMETRY OF GAMES AND THEIR SOLVERS

Anonymous authors
Paper under double-blind review

ABSTRACT

Understanding where learning algorithms converge, cycle, or otherwise change behaviour is fundamental to solving games in economics and multi-agent systems. Existing game taxonomies are largely derived from analytical structure, such as potential, zero-sum, or monotone regimes, and provide powerful guarantees in restricted settings, but their boundaries need not coincide with realised solver dynamics or predict finite-time behaviour. We introduce *behavioural cartography*, a low-dimensional representation of games organised directly by the learning dynamics they induce. Rather than assigning only a structural class, it identifies which games elicit similar solver behaviour, where algorithmic regimes change, and how solver competence varies across game space. We construct the atlas from mechanistically distinct primitives spanning gradient, optimistic/extragradient, and fictitious-play families, evaluated across solver settings and initialisations. Games with similar induced dynamics are placed nearby in an oracle atlas $z^\star(G)$, while a payoff-based model learns to locate unseen games in the same space, respecting strategic invariances including action relabelling and player exchange. Crucially, the atlas generalises beyond its construction: it predicts withheld properties such as recurrence and reversal and transfers to solver primitives excluded entirely from atlas construction. Across heterogeneous solvers, the learned atlas strongly preserves which games produce similar dynamics (pairwise-distance Spearman 0.917); when a solver is withheld, its behaviour can still be inferred from limited observations, reducing error by 30% relative to a constant reference at 205 labels. The resulting cartography reveals coherent solver-specific convergence and sensitivity fields, continuous behavioural variation along controlled structural paths, and distinctions not fully captured by conventional game classifications. Together, these results provide a predictive empirical atlas of game solving that complements analytical taxonomy by showing how algorithms behave across game space, how their capabilities relate, and how newly introduced solvers can be characterised.

1 INTRODUCTION

For many applications of game theory, the practical question is not only what equilibrium a game possesses, but how a learning algorithm will behave while attempting to solve it. The same solver may converge rapidly on one game, cycle on another, or depend strongly on its initial condition. Understanding this variation is therefore key to choosing, analysing, and designing learning algorithms.

Classical game theory provides powerful structural tools for parts of this problem. Potential structure supports descent-style arguments, rotational and zero-sum structure motivates extra-gradient and optimistic methods, and specialised results characterise response-based and monotone dynamics Balduzzi et al. (2018); Daskalakis & Panageas (2018); Antonakopoulos et al. (2020); Leslie et al. (2020). These guarantees are precise under identifiable assumptions, but their organising principle is analytical structure rather than realised learning behaviour. The issue is not that classes such as potential or zero-sum games lack crisp boundaries, but that those boundaries need not coincide with where an algorithm changes behaviour. Outside known sufficient conditions, they therefore provide limited guidance about whether a particular solver will converge, cycle, or become unstable.

Figure 1 illustrates the discrepancy between structural classification and realised learning behaviour. For GDA, games belonging to the same structural family can exhibit markedly different convergence outcomes, while structurally distinct games may induce similar behaviour. This observation motivates

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

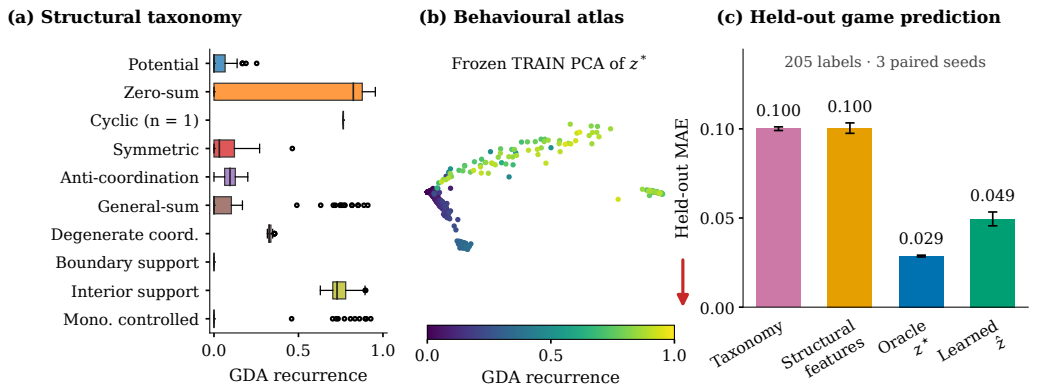


Figure 1: **From structural taxonomy to behavioural cartography.** (a) Distribution of GDA recurrence within each structural game family, revealing substantial within-family behavioural variation. (b) The same games projected onto the frozen TRAIN-fitted PCA of the oracle behavioural atlas, coloured by measured GDA recurrence, which is excluded from atlas construction. The two-dimensional projection is used solely for visualisation. (c) Held-out recurrence prediction using taxonomy labels, structural features, oracle coordinates $z^\star$, and learned coordinates $\hat{z}$. Bars show mean MAE over three paired label seeds at 205 labels, with error bars indicating sample standard deviation. Lower is better.

organising games according to their induced solver dynamics, providing an empirical perspective complementary to conventional structural taxonomies.

This creates a gap between the taxonomy of games and the behaviour of the algorithms used to solve them. Games that appear similar under conventional descriptors may induce very different dynamics, while games from different analytical classes may elicit similar solver responses. Terminal measures such as final exploitability further obscure how an outcome was reached. What is missing is a representation in which proximity has an operational meaning: games should be close because learning algorithms actually behave similarly on them.

We introduce *behavioural cartography*, a low-dimensional representation organised directly by the learning dynamics games induce. Rather than merely assigning a game to an a priori class, the resulting atlas characterises which games elicit similar solver behaviour, where dynamical regimes change, and how solver competence varies across game space. In this sense, it provides an empirical complement to analytical taxonomy. To construct the atlas, we use mechanistically distinct learning primitives spanning major families of learning-in-games algorithms, including gradient-based, mirror/proximal, optimistic and extra-gradient, and fictitious-play mechanisms. Each is evaluated across prescribed solver configurations and initialisations. Time-indexed measures of progress towards equilibrium and strategy movement define a behavioural distance between games, which is embedded into a compact oracle atlas $z^\star(G)$. A payoff-based recogniser then learns to place an unseen game directly into the same coordinate system without rerunning the full solver panel. The construction respects strategically irrelevant transformations such as action relabelling and player exchange, while aggregating variation across solver settings and initialisations.

A useful representation should also generalise beyond the observations used to construct it. We therefore test two forms of transfer. First, we predict behavioural quantities excluded from the source geometry, including recurrence and reversal. Second, in strict leave-one-solver-out experiments, the target solver is removed from the behavioural distance, embedding, and representation training, and is only subsequently observed through a limited number of labelled examples. Successful transfer means that structure learned from existing solver mechanisms can help characterise a newly introduced solver without reconstructing game space from scratch. The shared coordinates also expose relationships between solvers themselves: regions in which their behavioural fields align reveal common structure, while divergence identifies mechanism-specific responses.

Our experiments support this picture. Across heterogeneous solvers, the learned atlas strongly preserves which games induce similar learning dynamics, achieving pairwise-distance Spearman 0.917 against similarities measured directly from solver trajectories. It also predicts behavioural properties withheld from construction. Most notably, when a solver is excluded entirely from atlas

construction, its behaviour can still be inferred from limited observations: at 205 labels, the learned atlas reduces error by approximately 30% relative to a constant reference. These results indicate that the representation captures reusable structure shared across solver mechanisms rather than merely reproducing the observations used to build it.

The cartographic view goes beyond prediction. Behavioural quantities and solver performance become fields over a common game space, allowing us to examine where algorithms converge, exhibit persistent movement, or become sensitive to solver configurations and initialisations. Controlled paths through game space further trace how learning behaviour evolves under systematic changes in game structure. Our cartographic analysis reveals coherent yet solver-dependent convergence and sensitivity fields, together with systematic behavioural variation along controlled structural paths. These findings provide an empirical perspective on algorithmic regimes that complements conventional structural classifications.

Our contributions are threefold. **(i) Behavioural cartography of games.** We introduce an empirical geometry in which similarity is defined by induced learning dynamics across mechanistically distinct solver primitives. **(ii) A transferable representation of solver behaviour.** We recover this atlas from game payoffs while preserving strategic invariances, and show transfer to behavioural properties and solver mechanisms excluded from construction, including strict leave-one-solver-out prediction. **(iii) Empirical maps of algorithmic regimes.** We use the atlas to characterise solver-dependent behavioural fields and their variation along controlled structural paths, revealing behavioural differences within structural game families and similarities across families that are not fully captured by conventional game taxonomies.

Together, these results complement a taxonomy of what games *are* with an empirical account of how learning algorithms *behave* on them, providing a common framework for understanding existing solvers and rapidly characterising new ones.

2 Related Work

Game structure and learning dynamics. A substantial literature relates game structure to learning behaviour. Potential games admit a global potential function Monderer & Shapley (1996), while finite games can be decomposed into potential, harmonic, and nonstrategic components Candogan et al. (2011). Related decompositions of differentiable games separate potential-like and rotational structure Balduzzi et al. (2018), underpinning convergence analyses for extra-gradient, optimistic, and response-based methods under particular structural conditions Daskalakis & Panageas (2018); Antonakopoulos et al. (2020); Mokhtari et al. (2020); Leslie et al. (2020); Swenson et al. (2018). At the same time, learning dynamics can exhibit periodicity, recurrence, spurious convergence, and dependence on initialisation or iterate averaging Vlatakis-Gkaragkounis et al. (2019); Golowich et al. (2020). These results provide precise accounts of particular mechanisms, but not a common representation of finite-time behaviour across heterogeneous games and solvers.

Solver performance and behavioural representations. Per-instance algorithm selection uses problem characteristics to predict which algorithms will perform well Kerschke et al. (2019), demonstrating that algorithmic performance can reveal meaningful instance structure. Behaviour-based representations have similarly been studied in reinforcement learning, where policy-similarity metrics define geometries that support transfer across environments Agarwal et al. (2021). Our setting instead uses the joint responses of multiple game-solving mechanisms to define similarity between games, retaining variation across trajectories, solver configurations, and initialisations rather than reducing an instance to an algorithm ranking or terminal score.

Representation learning for games. Recent architectures explicitly respect the symmetries of normal-form games. NfgTransformer, for example, incorporates action-permutation equivariance and supports tasks including equilibrium prediction and game ranking Liu et al. (2024). Our question is complementary: *what geometry should organise a game representation when the object of interest is learning behaviour?* We define that geometry empirically from solver responses and learn to recover its coordinates directly from game payoffs.

These strands provide structural theories of games, convergence analyses for particular algorithms, methods for algorithm selection, and symmetry-aware game representations. What remains missing

is a shared, behaviour-grounded representation that spans solver mechanisms, can be inferred from the game itself, and remains informative for behaviours and solvers not used to construct it.

3 BEHAVIOURAL CARTOGRAPHY

Our aim is to represent games according to the learning behaviour they induce rather than only through an a priori structural class. We separate the construction into three objects,

$$G \longrightarrow d_{\mathrm{beh}} \longrightarrow z^\star(G) \longrightarrow \hat{z}(G),$$

where d_{beh} is defined directly from measured solver responses, $z^\star$ is a low-dimensional oracle representation of that empirical distance, and $\hat{z}$ places a game into the same coordinate system from its payoffs alone. Keeping these stages distinct lets us identify whether predictive or cartographic failures arise from the behavioural notion of similarity, from low-dimensional compression, or from recovering the atlas from payoffs.

Games and induced solver responses. Let $\mathcal{G}$ denote a collection of two-player normal-form games $G = (A, B)$, with $A, B \in \mathbb{R}^{n\times n}$, and let $\mathcal{S}$ be a heterogeneous library of learning algorithms. For solver $s \in \mathcal{S}$, configuration $h \in H_s$, and initial joint strategy $\theta^{(k)}$, we observe a finite-horizon trajectory

$$\tau_{G,s,h,k} = (\theta_{G,s,h,k,t})_{t=0}^{T}, \qquad \theta_{G,s,h,k,0} = \theta^{(k)}. \tag{1}$$

Every game is evaluated using the same prescribed configuration and initialisation panels. The solver definitions, parameter panels, and numerical protocol are given in Appendix A.2.

Solver library and evaluation panel. The atlas is built from seven mechanistically distinct learning primitives: gradient descent–ascent, entropic mirror descent, a damped proximal update, averaging/anchoring, extragradient, optimistic gradient, and fictitious play. Across the library we evaluate 31 prescribed solver configurations, using a common panel of 19 initialisations and a horizon of 200 updates. The same panel is applied to every game, so behavioural proximity reflects systematic differences between games rather than changes in the evaluation protocol. Full update rules and configuration values are given in Appendix A.2.

Behavioural observations. We construct the atlas from two complementary time-indexed signals: exploitability, which measures progress towards equilibrium, and strategy movement, which measures the magnitude of successive updates. Writing $\theta_t = (x_t, y_t)$, these are

$$r^{(\mathrm{gap})}_{G,s,h,k,t} = e_G(\theta_t), \qquad r^{(\mathrm{mov})}_{G,s,h,k,t} = \frac{1}{2}\left(\|x_t - x_{t-1}\|_1 + \|y_t - y_{t-1}\|_1\right). \tag{2}$$

The full exploitability definition and scaling procedure are given in Appendix B.1. Crucially, recurrence, reversal, and related higher-order trajectory properties are *not* used to construct the atlas; they are retained as held-out quantities for testing whether the representation captures structure beyond its source observations. For each solver s, time t, and source signal $b \in \mathcal{B}_{\mathrm{src}} = \{\mathrm{gap}, \mathrm{mov}\}$, we pool the normalised responses over the prescribed configurations and initialisations:

$$\mu^{(b)}_{G,s,t} = \frac{1}{N_s}\sum_{h\in H_s}\sum_{k=1}^{K} \delta_{\tilde{r}^{(b)}_{G,s,h,k,t}}, \qquad N_s = |H_s|K. \tag{3}$$

Thus each solver contributes a distribution of responses over its prescribed configurations and initialisations rather than a single privileged run. The same evaluation panel is applied to every game, so differences in the resulting atlas reflect game-dependent solver behaviour rather than changes in the measurement protocol.

Behavioural distance. We compare two games by the Wasserstein discrepancy between these response distributions:

$$d_{\mathrm{beh}}(G, G') = \frac{1}{|\mathcal{S}|T|\mathcal{B}_{\mathrm{src}}|}\sum_{s\in\mathcal{S}}\sum_{t=1}^{T}\sum_{b\in\mathcal{B}_{\mathrm{src}}} W_1\left(\mu^{(b)}_{G,s,t}, \mu^{(b)}_{G',s,t}\right). \tag{4}$$

Solver–signal responses are normalised using robust TRAIN-derived scales, after which each solver family and source signal receives equal aggregate weight. The resulting distance compares the

time-indexed distributions of progress and movement induced by the solver panel. It intentionally discards run identity, configuration correspondence, and complete cross-time trajectory structure; the exact empirical W_1 expression and these design choices are detailed in Appendix B.1. This separation is useful experimentally because higher-order properties such as recurrence and reversal can then be treated as genuinely held-out behavioural tests.

Low-dimensional oracle atlas. The empirical distance gives a behavioural reference geometry over games. We use classical multidimensional scaling on TRAIN landmarks to obtain

$$z^\star : \mathcal{G} \to \mathbb{R}^d,$$

with Euclidean distances in $z^\star$ approximating d_{beh}. The landmark set, normalisation, and MDS eigensystem are fixed using TRAIN, and unseen games are placed using the corresponding out-of-sample extension. The full construction and dimensionality-selection protocol are given in Appendix B.3. This oracle atlas is the object used to test whether the measured solver-response structure is both compact and cartographically useful.

Payoff-only recovery. Direct construction of $z^\star(G)$ requires running the solver panel. To make the atlas usable for a new game, we learn a recogniser

$$\hat{z}(G) = f_\phi(A, B) \approx z^\star(G), \qquad \mathcal{L}_{\mathrm{coord}}(\phi) = \frac{1}{|\mathcal{D}_{\mathrm{rec}}|} \sum_{G \in \mathcal{D}_{\mathrm{rec}}} \|f_\phi(G) - z^\star(G)\|_2^2 . \tag{5}$$

The recogniser is invariant to independent action relabellings and exchange of the two players, and downstream behavioural labels are excluded from its supervision. Across the full transformation audit, the maximum coordinate discrepancy is 6.9×10^{-5}, while the evaluated downstream predictions and neighbour selections are unchanged. Architecture, optimisation, and the complete invariance audit are reported in Appendices A.4 and C.3. **Generalisation beyond the source atlas.** The representation is useful only if it captures more than the signals from which it was built. We therefore freeze $z^\star$ and $\hat{z}$ and evaluate lightweight readouts for additional behavioural quantities. In strict leave-one-solver-out evaluation, the target solver is removed from the source distance, MDS construction, and recogniser supervision before any of its behavioural labels are used. This asks whether structure induced by existing solver mechanisms can help characterise a solver that played no role in constructing the atlas. Section 4.3 reports transfer to both withheld behavioural quantities and excluded solvers.

From representation to cartography. Finally, we treat measured behavioural quantities and solver performance as fields over the shared game space. Scientific neighbourhoods, competence regions, and boundaries are computed in the full d-dimensional atlas; two-dimensional projections are used only for visualisation. Regime boundaries arise from level-set or relative-performance crossings of these fields and therefore do not require the atlas itself to be discontinuous. Controlled structural paths provide a complementary test of whether movement through the atlas tracks measured changes in solver behaviour. The complete protocol is given in Appendix C.4, with the principal empirical results in Section 4.4.

4 EXPERIMENTS

Our experiments follow the construction of the atlas. We first ask whether heterogeneous solvers share game-level behavioural structure; we then measure how much of that structure survives low-dimensional compression and payoff-only recovery; next we test transfer to withheld behaviours and solvers; finally, we examine whether the resulting space supports a meaningful cartography of algorithmic regimes.

Games and evaluation protocol. We use a coverage-oriented corpus of $35{,}804$ two-player normal-form games spanning canonical and interpolated structural regimes. TRAIN, DEV, and TEST are partitioned by generative root so that variants of the same underlying game do not cross splits. The principal behavioural evaluation uses $2{,}048$ TRAIN and 512 DEV games. Behavioural-distance calibration, landmark selection, and the MDS eigensystem use TRAIN only; representation and dimensionality choices use TRAIN and DEV. TEST is not used for representation or model selection; the results reported below follow the stated TRAIN/DEV evaluation protocol unless otherwise specified. Dataset construction, solver configurations, and the initialisation protocol are detailed in Appendices A.1 and A.2.

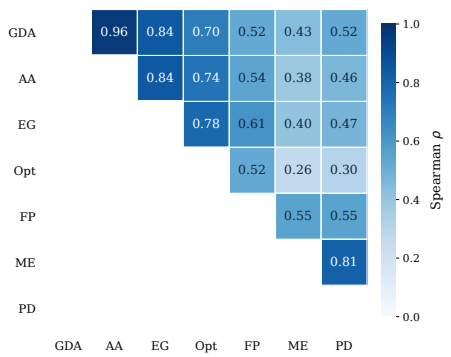


Figure 2: **Cross-solver agreement in behavioural game distances.** Pairwise Spearman correlations between solver-specific behavioural distances computed over the same game population. Each entry measures the agreement in game-pair distance rankings induced by two solvers. The correlations reveal substantial shared organisation across learning mechanisms, alongside solver-dependent differences.

Representations and baselines. We evaluate the raw behavioural distance d_{beh}, its oracle embedding $z^\star$, and the payoff-derived coordinates $\hat{z}$. These isolate the information in the measured solver responses, the effect of low-dimensional compression, and the additional approximation incurred by payoff-only recognition. Baselines include direct payoff prediction, NfgTransformer, response-graph representations, Candogan-based structural features, and behaviour-pretrained controls where permitted by the evaluation protocol. Where a baseline uses source behavioural supervision, we match the permissible source information, target-label subsets, and downstream readout protocol.

Readouts and metrics. Frozen representations are evaluated with both k-nearest-neighbour and Ridge readouts, probing local behavioural similarity and linear accessibility, respectively. Unless stated, target-label budgets are 20, 102, and 205 games. We report MAE and normalised MAE,

$$\mathrm{NMAE}(\widehat{y}; y) = \frac{\mathrm{MAE}(\widehat{y}, y)}{\mathrm{MAE}(\widetilde{y}_{\mathrm{train}}, y)},$$

where $\widetilde{y}_{\mathrm{train}}$ is the constant-median prediction fitted from the available labelled subset. Thus NMAE $= 1$ corresponds to the constant reference and lower is better. Rare or nearly constant targets use a common TRAIN-defined non-degeneracy criterion; complete target definitions are in the appendix.

4.1 Do Different Solvers Share a Common Behavioural Organisation?

Behavioural cartography is useful only if distinct learning mechanisms share some game-level organisation. Figure 2 shows substantial agreement between several solver-specific behavioural distances, together with clear solver-dependent variation: the same games tend to be behaviourally related across multiple mechanisms, but not identically so. This suggests that different learning mechanisms induce overlapping behavioural organisations of game space, providing an empirical basis for constructing a shared atlas.

A stronger test removes an entire solver from the source representation. For each target solver s, we reconstruct the behavioural distance using only $\mathcal{S} \setminus \{s\}$ and ask whether the resulting source-only structure remains informative about s. At 205 labelled games, the raw source-only distance achieves macro-average NMAE 0.626 across the seven excluded solvers, with every solver below the constant reference. Behavioural similarity induced by one collection of mechanisms therefore contains information about a mechanism that was not used to define it. The oracle and learned versions of this transfer test are examined in Section 4.3.

4.2 How Much Structure Survives Compression and Payoff-Only Recovery?

We next separate two sources of information loss. The oracle atlas preserves the broad distance ordering of the empirical behavioural geometry, while payoff-only recognition introduces additional distortion. On DEV, the learned atlas achieves pairwise-distance Spearman 0.917 against the raw behavioural distance (Appendix B.4). A dimensionality ablation further shows that the oracle retains

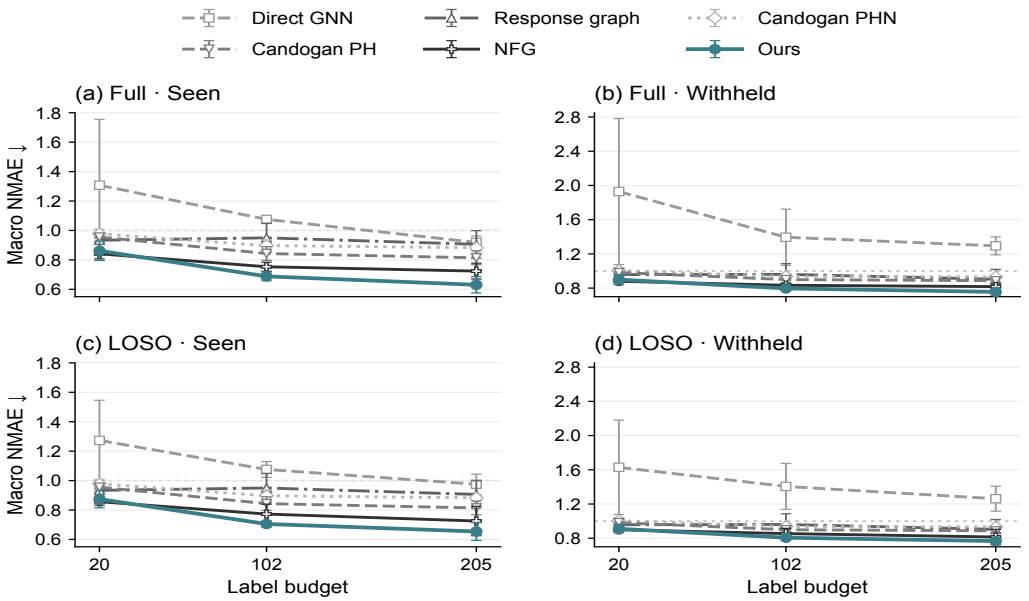


Figure 3: **Behavioural transfer across label budgets and solver mechanisms.** Macro-average NMAE (lower is better) under full-solver and leave-one-solver-out (LOSO) evaluation. Seen quantities are included in the behavioural distance, whereas withheld quantities (recurrence and reversal) are excluded from its construction. All methods use a Ridge readout on frozen representations with identical labelled subsets at budgets of 20, 102, and 205 games. Points show the mean over three paired label seeds, with error bars indicating sample standard deviation.

substantial geometric fidelity even at four dimensions (Spearman 0.981); increasing the dimension beyond 12 yields only modest geometric improvements without corresponding gains in strict LOSO prediction (Appendix B.5). These results support the use of a compact behavioural representation.

4.3 Does the Atlas Generalise Beyond the Behaviours and Solvers Used to Construct It?

A reusable behavioural representation should remain informative beyond the responses used to build it. We test two complementary forms of transfer. *Behaviour transfer* predicts trajectory properties withheld from distance construction, including recurrence and reversal. *Solver transfer* uses a strict leave-one-solver-out (LOSO) protocol in which the target solver is excluded from distance construction, embedding, and recogniser training. Their combination asks whether the atlas can predict a withheld property of a solver that played no role in constructing it.

Useful information is accessible with limited supervision. Figure 3 reports the endpoint supervision regimes, with the complete three-budget comparison in Appendix C.1. The atlas also performs strongly under a simple KNN readout. The corresponding Ridge comparison is reported in the appendix. Together, these results show that the encoded behavioural information is accessible both locally and through a simple linear map.

The atlas predicts behaviours it was not built from. The same representation supports prediction of trajectory properties excluded from the source distance, including recurrence and reversal. This is a stronger test than reconstructing exploitability or movement: success requires the atlas to capture structure that covaries with higher-order properties of the trajectories rather than merely reproducing its source measurements.

Behavioural structure transfers to an excluded solver. Strict LOSO provides the strongest cross-mechanism test. No responses from the target solver participate in the distance, oracle embedding, or learned representation. With a kNN readout, the learned atlas achieves overall NMAE 0.889, 0.745, and 0.699 at 20, 102, and 205 labels, respectively. At the largest budget this is approximately 30% lower error than the constant reference. Thus a solver absent from atlas construction can still be characterised from relatively few subsequent observations. The joint solver–behaviour setting strengthens this result by withholding both the mechanism and the target behavioural quantity. Full per-target, per-solver, oracle, Ridge, and nonlinear-readout results are reported in Appendix C.1.

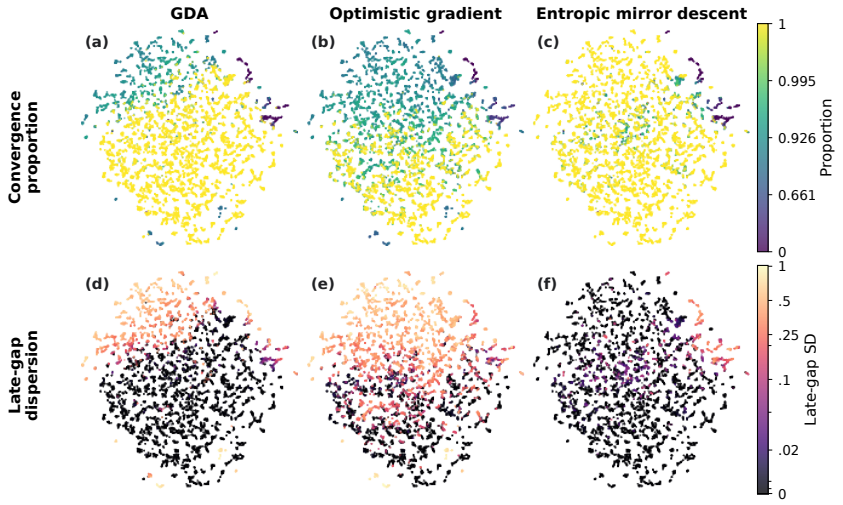


Figure 4: **Solver-specific behavioural fields over a shared game representation.** Each solver is evaluated on the same 20,512 games. The left and right maps within each panel show measured convergence proportion and late-stage exploitability dispersion, respectively. All panels share a fixed two-dimensional projection of the behavioural atlas. Colours indicate measured responses, and all quantitative geometric analyses use the full 12-dimensional representation.

Effect of behavioural pretraining. To examine whether transfer gains arise from behavioural pretraining alone, we compare the learned atlas with a matched payoff encoder pretrained to regress source-behaviour summaries. Under strict LOSO evaluation, the learned atlas achieves NMAE 0.699 at 205 labels, compared with 0.803 for the ordinary pretrained encoder, with lower error across all seven excluded solvers. This result supports the utility of organising source responses through a shared behavioural geometry (Appendix C.2).

Finding. The behavioural atlas is not merely a compressed record of the solver panel. Its organisation transfers to withheld trajectory properties and to mechanisms excluded from construction, supporting rapid characterisation of newly introduced solvers from limited behavioural evidence.

4.4 Solver-Specific Behavioural Cartography

We examine whether the shared behavioural atlas provides a spatial organisation of solver-dependent learning behaviour and retains information about behavioural variation under controlled changes in game structure. We evaluate measured convergence and sensitivity fields over a common representation, followed by behavioural readouts along controlled structural paths. All quantitative geometric analyses use the full 12-dimensional representation; two-dimensional projections are used only for visualisation.

Behavioural fields and local organisation. Figure 4 visualises convergence proportion and late-stage exploitability dispersion across the same 20,512 games for three representative solvers. Convergence proportion measures the fraction of prescribed configurations and initialisations satisfying the empirical convergence criterion, while dispersion measures variation in late-stage exploitability across the evaluation panel. The resulting fields exhibit common spatial organisation alongside distinct solver-dependent patterns. GDA, optimistic gradient, and entropic mirror descent achieve full convergence across the prescribed evaluation panel on 71.8%, 35.3%, and 83.3% of games, respectively. Their sensitivity fields also exhibit different spatial distributions, illustrating how the shared atlas accommodates heterogeneous algorithmic responses.

Controlled structural paths. We further examine how controlled changes in game structure are reflected in the atlas and its associated behavioural responses. We consider a path interpolating between zero-sum and common-interest games. Given fixed, centred and normalised payoff matrices Z and C, the two players' payoffs are defined as

$$A_\lambda = (1-\lambda)Z + \lambda C, \qquad B_\lambda = -(1-\lambda)Z + \lambda C, \quad \lambda \in [0,1]. \tag{6}$$

The endpoints correspond to a zero-sum game and a common-interest game, respectively. The interpolation progressively changes their relative contributions.

Fig. 5 follows eleven games along this path and examines the resulting GDA recurrence, a withheld dynamical quantity measuring repeated visits to previously encountered strategy states. The games

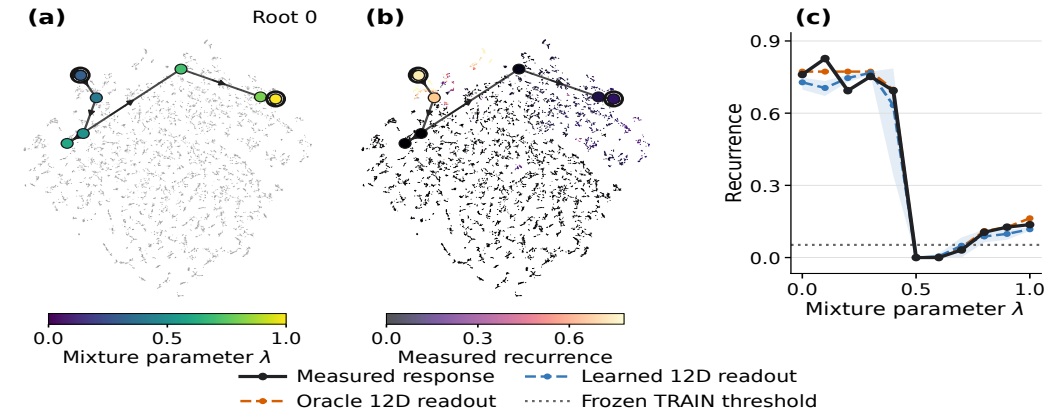


Figure 5: **Behavioural variation along a controlled game-mixture path.** (a) Eleven games interpolating between zero-sum ($\lambda = 0$) and common-interest ($\lambda = 1$) payoffs, shown in a fixed two-dimensional projection of the shared behavioural atlas. Colours indicate the mixture parameter. (b) The same path coloured by measured GDA recurrence, with background games coloured by their corresponding behavioural responses. (c) Measured recurrence compared with readouts from the oracle and learned 12-dimensional representations. Shading indicates the range of learned predictions across three paired label seeds, and the dotted line denotes the recurrence threshold fixed on TRAIN. All behavioural readouts use the full 12-dimensional representations; the 2-dimensional projection is used only for visualisation. Complete results for the prescribed controlled paths are in Appendix C.4.

trace an ordered path through the shared atlas, accompanied by substantial variation in measured recurrence. Recurrence decreases sharply between $\lambda = 0.4$ and 0.5, before increasing gradually towards the common-interest endpoint. Both oracle and learned atlas readouts recover this overall progression, including the sharp behavioural change between adjacent path points and the subsequent crossing of the frozen TRAIN threshold.

Finding. The shared atlas organises heterogeneous solver responses into coherent behavioural fields and captures systematic variation in withheld dynamics along a controlled game-mixture path. Oracle coordinates closely preserve the measured behavioural progression, while payoff-only recognition recovers its principal behavioural changes with measurable local error.

5 Discussion and Conclusion

Our results support a shared behavioural organisation of game space. Structure induced by one collection of learning mechanisms remains informative about others, survives low-dimensional compression, and can be recovered in broad form from game payoffs. The representation also generalises beyond its construction, supporting prediction of withheld dynamical properties and of solvers excluded from the source atlas. The cartographic analysis provides the complementary test of whether this organisation also identifies where distinct algorithmic regimes arise. We observe coherent yet solver-dependent convergence and sensitivity fields, with behavioural variation along controlled structural paths that is substantially preserved by the oracle atlas and partially recovered from game payoffs. These findings reveal an empirical organisation of learning dynamics that extends beyond conventional structural game classifications. Additional cross-size experiments show that the frozen payoff recogniser transfers from 3×3 games to 4×4 and 5×5 games without target-size behavioural labels or encoder retraining(Appendix C.6).

The atlas is deliberately protocol-conditioned rather than canonical: it reflects the chosen solver library, configurations, initialisations, horizon, and behavioural signals. Moreover, the current distance preserves time-indexed distributions of progress and movement rather than complete run-wise trajectories. The decomposition $d_{\mathrm{beh}} \to z^\star \to \hat{z}$ therefore remains important empirically, separating limitations of the behavioural distance from compression loss and payoff-only recognition; in the present results, the latter is most visible in fine local neighbourhoods. More broadly, behavioural cartography complements analytical taxonomy with an empirical account of how learning algorithms respond across game space, offering a common substrate for comparing existing mechanisms and rapidly characterising newly introduced solvers from limited behavioural evidence.

AI USE STATEMENT

We used generative artificial intelligence (AI) tools for language editing, drafting support, literature search, and to support some implementation and analysis methods. The authors reviewed all AI-assisted text, technical claims, mathematical formulations, experimental specifications, analyses, and code, and take full responsibility for the final content.

REPRODUCIBILITY STATEMENT

The paper specifies the game corpus and data splits, payoff normalisation, solver library and configuration panels, initialisation protocol, behavioural targets, behavioural distance, oracle MDS construction, payoff-based recogniser, downstream readouts, baselines, leave-one-solver-out evaluation, and cartographic analyses. Appendices A–C provide the full experimental and implementation details, including solver-response construction, behavioural-distance computation, representation learning, invariance checks, transfer protocols, and controlled structural-path experiments.

REFERENCES

Rishabh Agarwal, Marlos C Machado, Pablo Samuel Castro, and Marc G Bellemare. Contrastive behavioral similarity embeddings for generalization in reinforcement learning. *arXiv preprint arXiv:2101.05265*, 2021.

Kimon Antonakopoulos, E Veronica Belmega, and Panayotis Mertikopoulos. Adaptive extra-gradient methods for min-max optimization and games. *arXiv preprint arXiv:2010.12100*, 2020.

David Balduzzi, Sebastien Racaniere, James Martens, Jakob Foerster, Karl Tuyls, and Thore Graepel. The mechanics of n-player differentiable games. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 354–363. PMLR, 2018. URL https://proceedings.mlr.press/v80/balduzzi18a.html.

Ozan Candogan, Ishai Menache, Asuman Ozdaglar, and Pablo A. Parrilo. Flows and decompositions of games: Harmonic and potential games. *Mathematics of Operations Research*, 36(3):474–503, 2011. doi: 10.1287/moor.1110.0500. URL https://doi.org/10.1287/moor.1110.0500.

Constantinos Daskalakis and Ioannis Panageas. The limit points of (optimistic) gradient descent in min-max optimization. *Advances in neural information processing systems*, 31, 2018.

Noah Golowich, Sarath Pattathil, Constantinos Daskalakis, and Asuman Ozdaglar. Last iterate is slower than averaged iterate in smooth convex-concave saddle point problems. In *Proceedings of the Thirty-Third Conference on Learning Theory*, volume 125 of *Proceedings of Machine Learning Research*, pp. 1758–1784. PMLR, 2020.

Pascal Kerschke, Holger H. Hoos, Frank Neumann, and Heike Trautmann. Automated algorithm selection: Survey and perspectives. *Evol. Comput.*, 27(1):3–45, March 2019. ISSN 1063-6560. doi: 10.1162/evco_a_00242. URL https://doi.org/10.1162/evco_a_00242.

David S Leslie, Steven Perkins, and Zibo Xu. Best-response dynamics in zero-sum stochastic games. *Journal of Economic Theory*, 189:105095, 2020.

Siqi Liu, Luke Marris, Georgios Piliouras, Ian Gemp, and Nicolas Heess. NfgTransformer: Equivariant Representation Learning for Normal-Form Games. In *International Conference on Learning Representations*, 2024.

Aryan Mokhtari, Asuman Ozdaglar, and Sarath Pattathil. A unified analysis of extra-gradient and optimistic gradient methods for saddle point problems: Proximal point approach. In *International Conference on Artificial Intelligence and Statistics*, pp. 1497–1507. PMLR, 2020.

Dov Monderer and Lloyd S. Shapley. Potential games. *Games and Economic Behavior*, 14(1): 124–143, 1996. doi: 10.1006/game.1996.0044. URL https://doi.org/10.1006/game.1996.0044.

Brian Swenson, Ryan Murray, and Soummya Kar. On best-response dynamics in potential games. *SIAM Journal on Control and Optimization*, 56(4):2734–2767, 2018.

Emmanouil-Vasileios Vlatakis-Gkaragkounis, Lampros Flokas, and Georgios Piliouras. Poincaré recurrence, cycles and spurious equilibria in gradient-descent-ascent for non-convex non-concave zero-sum games. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://proceedings.neurips.cc/paper_files/paper/2019/file/6c7cd904122e623ce625613d6af337c4-Paper.pdf.

A Experimental Setup and Implementation Details

This appendix describes the game corpus, solver-response protocol, and implementation of the payoff-based recogniser. All representation construction and model selection follow the TRAIN/DEV partition, with the corresponding experimental settings held fixed for downstream evaluation.

A.1 Dataset Construction

Game corpus. We use a coverage-oriented corpus of $35{,}804$ two-player normal-form games generated from canonical structural families and interpolations between them. The corpus spans potential-like, harmonic, zero-sum, symmetric, and mixed game structures. Candidate games are sampled to improve coverage across structural characteristics, followed by deduplication. This construction provides a heterogeneous collection for examining solver behaviour across game space. The reported results characterise the resulting evaluation distribution and do not assume a natural probability distribution over games.

Payoff normalisation. For the original payoff matrices $\widetilde{A}, \widetilde{B} \in \mathbb{R}^{n\times n}$, we first centre each player's payoffs:

$$\begin{aligned} \bar{a} = \frac{1}{n^2}\sum_{i,j}\widetilde{A}_{ij}, \qquad \bar{b} = \frac{1}{n^2}\sum_{i,j}\widetilde{B}_{ij}, \\ A^\circ = \widetilde{A} - \bar{a}\mathbf{1}\mathbf{1}^\top, \quad B^\circ = \widetilde{B} - \bar{b}\mathbf{1}\mathbf{1}^\top. \end{aligned} \tag{7}$$

A common scale is subsequently applied to both players:

$$\begin{aligned} c_G = \max\left\{\max_{i,j}|A^\circ_{ij}|, \max_{i,j}|B^\circ_{ij}|, \varepsilon_{\text{pay}}\right\}, \\ A = \frac{A^\circ}{c_G}, \qquad B = \frac{B^\circ}{c_G}, \end{aligned} \tag{8}$$

where $\varepsilon_{\text{pay}} > 0$ is a numerical floor. The shared scaling preserves the relative payoff magnitudes of the two players within each game.

Data partitioning. The corpus is partitioned into TRAIN, DEV, and TEST at the level of generative roots, ensuring that variants of the same underlying game remain within one partition. The principal behavioural evaluation uses $2{,}048$ TRAIN games and 512 DEV games. Behavioural-distance calibration, landmark construction, and the MDS eigensystem use TRAIN exclusively. The payoff recogniser is supervised using oracle coordinates from a larger TRAIN subset. Representation and model selection use TRAIN and DEV. TEST is not used for representation or model selection; all reported evaluations identify explicitly which partition is used.

A.2 Solver Library and Response Protocol

We consider seven learning primitives spanning gradient-based, mirror-based, stabilised, predictive, and response-based mechanisms. Table 1 summarises the solver library. Each primitive contributes a separate response block to the behavioural distance.

For each solver $s \in \mathcal{S}$, let H_s denote its prescribed configuration set. The full evaluation panel comprises 31 solver configurations across the seven primitives. Each configuration is evaluated using a common panel of $K = 19$ initial joint strategies over $T = 200$ updates. The same configurations and initialisations are applied to every game.

For game G, the resulting response collection is

$$\mathcal{R}(G) = \{\tau_{G,s,h,k} : s \in \mathcal{S},\ h \in H_s,\ k \in \{1,\dots,K\}\}, \tag{9}$$

where $\tau_{G,s,h,k}$ is the trajectory defined in Eq. 1. Each solver therefore contributes $N_s = |H_s|K$ trajectories per game.

The configuration panel is fixed within each solver family. Variation across prescribed configurations and initialisations is retained in the empirical response distributions, allowing the atlas to characterise a solver's behaviour across the evaluation panel.

Table 1: **Solver library.** The primitives span distinct learning mechanisms and are evaluated using their prescribed configuration panels.

Solver	Update mechanism
Gradient descent–ascent (GDA)	Simultaneous projected gradient updates
Entropic mirror descent	Multiplicative strategy updates
Damped proximal update	Damped projected updates
Averaging and anchoring	Updates with a running strategy anchor
Extragradient	Extrapolation followed by gradient correction
Optimistic gradient	Updates using current and previous gradients
Fictitious play	Best-response-based empirical averaging

A.3 Behavioural Targets and Evaluation Protocol

We distinguish the time-indexed source observations used to construct the behavioural distance from the scalar targets used to evaluate representation transfer and cartographic organisation. The source geometry is defined by exploitability and successive strategy movement. Downstream evaluation considers three summaries derived from these source signals and two higher-order trajectory properties withheld from distance construction. All measurements follow the prescribed solver configurations and initialisations, with a common horizon of $T = 200$ updates and $K = 19$ initialisations per configuration.

Source observations. For a game $G = (A, B)$ and joint strategy $\theta_t = (x_t, y_t)$, the exploitability gap is

$$\begin{aligned} e_G(\theta_t) = &\max_{x' \in \Delta_n} {x'}^\top A y_t - x_t^\top A y_t \\ &+ \max_{y' \in \Delta_n} x_t^\top B y' - x_t^\top B y_t. \end{aligned} \tag{10}$$

The gap measures the sum of the two players' unilateral improvement opportunities and vanishes at a Nash equilibrium. Successive strategy movement is measured by

$$m_t = \frac{1}{4} \left(\|x_t - x_{t-1}\|_1 + \|y_t - y_{t-1}\|_1 \right). \tag{11}$$

Both quantities are evaluated at every iteration. For each game, solver, and time step, the responses are pooled over the prescribed configuration–initialisation panel, giving the empirical measures in Eq. 3. These time-indexed marginal distributions define the behavioural distance in Eq. 4. Recurrence and reversal are excluded from this construction.

Run-level behavioural targets. Let $g_t = e_G(\theta_t)$ and m_t denote the gap and movement of an individual solver run. We define the three source-derived targets as

$$a(\tau) = \frac{1}{T} \sum_{t=1}^{T} g_t, \tag{12}$$

$$\ell_g(\tau) = \frac{1}{20} \sum_{t=T-19}^{T} g_t, \tag{13}$$

$$\ell_m(\tau) = \frac{1}{40} \sum_{t=T-39}^{T} m_t. \tag{14}$$

Here $a(\tau)$ is the normalised area under the gap trajectory, denoted gap AUC, while $\ell_g(\tau)$ and $\ell_m(\tau)$ summarise late-stage exploitability and movement. Lower values indicate smaller average gaps or smaller late-stage strategy changes, respectively. Small movement alone does not establish proximity to equilibrium.

We additionally evaluate two withheld trajectory properties over the final 80-step window. Late recurrence measures the frequency with which a trajectory returns near an earlier strategy state after a substantive intervening excursion. An eligible return requires a temporal separation of at least 10

Table 2: **Behavioural targets for downstream transfer.** All quantities are first computed on individual trajectories and subsequently averaged over the prescribed solver configuration–initialisation panel. Seen denotes a target derived from a source signal; withheld denotes a trajectory property excluded from behavioural-distance construction.

Category	Target	Run-level definition	Window
Seen	Gap AUC	Normalised gap integral	200 steps
Seen	Late gap	Mean exploitability gap	Last 20
Seen	Late movement	Mean strategy movement	Last 40
Withheld	Late recurrence	Frequency of eligible returns	Last 80
Withheld	Late reversal	Fraction of reversing valid updates	Last 80

steps, a joint-strategy return distance no greater than 0.02, and accumulated intervening movement of at least 0.05. The excursion requirement prevents nearly stationary trajectories from being counted as recurrent solely because successive states remain close.

Writing $\mathcal{T}_{\mathrm{late}}$ for the final 80-state evaluation window, $\rho(\theta_t, \theta_j)$ for the joint-strategy distance used by the response protocol, and $M(j,t)$ for the accumulated movement between two observations, the recurrence indicator is

$$R_t = \mathbf{1}\left\{ \begin{array}{l} \exists j \in \mathcal{T}_{\mathrm{late}},\ j \le t-10 : \\ \rho(\theta_t, \theta_j) \le 0.02, \\ M(j,t) \ge 0.05 \end{array} \right\}. \tag{15}$$

The late-recurrence target, denoted $r(\tau)$, is the empirical return frequency obtained by aggregating these indicators over the prescribed late-stage window. This criterion measures finite-horizon return behaviour and does not establish mathematical periodicity.

Late reversal measures changes in the direction of successive strategy updates. Let

$$u_t = \theta_t - \theta_{t-1}, \tag{16}$$

and let $\mathcal{V}_{\mathrm{late}}$ contain the adjacent update pairs in the final 80-update window for which both updates satisfy the numerical non-zero criterion of the response protocol. We define

$$v(\tau) = \frac{1}{|\mathcal{V}_{\mathrm{late}}|} \sum_{t \in \mathcal{V}_{\mathrm{late}}} \mathbf{1}\left\{ \langle u_t, u_{t-1} \rangle < 0 \right\}. \tag{17}$$

The prescribed numerical convention is used when no valid update pair is available. Higher reversal indicates more frequent changes in update direction, without necessarily implying a persistent cycle.

Table 2 summarises the five targets used in the main transfer experiments. The first three are summaries of the source observations, while recurrence and reversal measure properties of the run-wise trajectories that are absent from the time-indexed marginal distance.

Panel aggregation. The downstream target for a game G and solver s is obtained by averaging each run-level quantity over the complete prescribed evaluation panel. For any behavioural functional $b(\tau)$, we define

$$y_{s,b}(G) = \frac{1}{N_s} \sum_{h \in H_s} \sum_{k=1}^{K} b(\tau_{G,s,h,k}), \qquad N_s = |H_s| K. \tag{18}$$

Each configuration–initialisation pair contributes equally within its solver family. This panel-mean convention applies to all five targets in the main behavioural-transfer experiments and to recurrence and reversal along the controlled structural paths.

Behavioural and solver generalisation. Let

$$\begin{aligned} \mathcal{B}_{\mathrm{seen}} &= \{\mathrm{AUC}, \text{late gap}, \text{late movement}\}, \\ \mathcal{B}_{\mathrm{withheld}} &= \{\text{recurrence}, \text{reversal}\}, \\ \mathcal{B}_{\mathrm{eval}} &= \mathcal{B}_{\mathrm{seen}} \cup \mathcal{B}_{\mathrm{withheld}}. \end{aligned} \tag{19}$$

Table 3: **Behavioural and solver-transfer settings.** Full evaluation uses all seven solvers to construct the source atlas. Strict LOSO reconstructs the atlas from the six remaining solvers and evaluates only the excluded solver. Seen and withheld refer to the behavioural target definitions, independently of whether the target solver participated in representation construction.

Setting	Source solvers	Evaluation targets	Count
Full–seen	All seven	Three source-derived quantities for every solver	21
Full–withheld	All seven	Recurrence and reversal for every solver	14
LOSO–seen	Six, excluding target solver	Three source-derived quantities of the excluded solver	3 per fold
LOSO–withheld	Six, excluding target solver	Recurrence and reversal of the excluded solver	2 per fold

With seven solver primitives, the full evaluation comprises $7 \times 5 = 35$ solver–behaviour targets, consisting of 21 source-derived and 14 withheld targets. The four evaluation settings are specified in Table 3.

For each strict LOSO fold, all responses from the target solver are excluded from behavioural-distance construction, oracle embedding, and recogniser supervision. The frozen source representation is subsequently evaluated using limited labelled observations from that solver. The term seen in LOSO–seen therefore refers exclusively to the behavioural quantity: the target solver's trajectories remain excluded from representation construction.

The overall LOSO result aggregates all five behavioural targets across the seven excluded-solver folds. Accordingly, the overall 205-label NMAE of 0.699 reported in the main text describes joint transfer across source-derived and withheld quantities. Results restricted to recurrence and reversal are reported separately under LOSO–withheld.

Additional cartographic quantities. The structural-family visualisation and solver-specific behavioural maps additionally use two gap-derived quantities. First, we measure the empirical low-gap proportion,

$$q_s(G) = \frac{1}{N_s} \sum_{h \in H_s} \sum_{k=1}^{K} \mathbf{1} \left\{ \ell_g(\tau_{G,s,h,k}) \leq 0.01 \right\}. \tag{20}$$

This quantity records the fraction of prescribed runs whose mean gap over the final 20 observations satisfies the fixed threshold. It measures finite-horizon attainment of a low-gap condition and does not impose an additional requirement on strategy movement.

Second, we quantify late-gap dispersion across the same evaluation panel. Writing

$$\bar{\ell}_{g,s}(G) = \frac{1}{N_s} \sum_{h \in H_s} \sum_{k=1}^{K} \ell_g(\tau_{G,s,h,k}), \tag{21}$$

the dispersion is

$$\sigma_{g,s}(G) = \sqrt{\frac{1}{N_s} \sum_{h \in H_s} \sum_{k=1}^{K} \left[\ell_g(\tau_{G,s,h,k}) - \bar{\ell}_{g,s}(G) \right]^2}. \tag{22}$$

This statistic measures variation in late-stage exploitability across both solver configurations and initialisations. The two cartographic quantities are derived from the gap observations and are not additional withheld behaviours. They are used to visualise solver-dependent behavioural fields over the frozen atlas, while recurrence and reversal provide complementary tests of trajectory information excluded from the source distance.

A.4 Payoff Recogniser

The oracle representation requires empirical solver responses. We therefore train a payoff-based recogniser $f_\phi : \mathcal{G} \to \mathbb{R}^d$ to recover the frozen oracle coordinates:

$$\mathcal{L}_{\mathrm{coord}}(\phi) = \frac{1}{|\mathcal{D}_{\mathrm{rec}}|} \sum_{G \in \mathcal{D}_{\mathrm{rec}}} \|f_\phi(G) - z^\star(G)\|_2^2 , \tag{23}$$

where $\mathcal{D}_{\mathrm{rec}}$ contains the TRAIN games used for recogniser supervision.

The architecture represents each normal-form game as a bipartite action graph. Shared permutation-equivariant message passing is followed by symmetric aggregation and a player-exchange-invariant readout. These operations ensure that strategically equivalent presentations have identical outputs in exact arithmetic. The numerical behaviour of the frozen recogniser is examined in Appendix C.3.

The recogniser is selected using DEV coordinate error. Downstream behavioural targets are excluded from its training objective, and the oracle coordinate system remains fixed throughout recogniser training and evaluation.

B Behavioural Distance and Oracle Construction

This appendix gives the construction underlying the behavioural distance and its low-dimensional representation. The source geometry is defined from time-indexed distributions of exploitability and strategy movement, aggregated across the prescribed solver configurations and initialisations.

B.1 Behavioural Observations and Normalisation

Let $G = (A, B)$ be a two-player normal-form game and let $\theta = (x, y)$ denote the joint mixed strategy, where $x, y \in \Delta_n$ and

$$\Delta_n = \left\{ v \in \mathbb{R}^n_{\geq 0} : \mathbf{1}^\top v = 1 \right\}. \tag{24}$$

The exploitability of θ is defined as

$$e_G(\theta) = \max_{x' \in \Delta_n} {x'}^\top A y - x^\top A y + \max_{y' \in \Delta_n} x^\top B y' - x^\top B y. \tag{25}$$

For each trajectory $\tau_{G,s,h,k}$, we extract two source observations:

$$\begin{aligned} r^{(\mathrm{gap})}_{G,s,h,k,t} &= e_G(\theta_{G,s,h,k,t}), \\ r^{(\mathrm{mov})}_{G,s,h,k,t} &= \frac{1}{2} \left(\|x_{G,s,h,k,t} - x_{G,s,h,k,t-1}\|_1 + \|y_{G,s,h,k,t} - y_{G,s,h,k,t-1}\|_1 \right). \end{aligned} \tag{26}$$

Both signals are evaluated for $t = 1, \ldots, T$. Exploitability measures progress towards equilibrium, while movement measures the magnitude of successive strategy updates.

To account for differences in numerical scale across solver families and behavioural signals, a positive robust scale $a_{s,b}$ is estimated from TRAIN responses for every $(s, b) \in \mathcal{S} \times \mathcal{B}_{\mathrm{src}}$, where $\mathcal{B}_{\mathrm{src}} = \{\mathrm{gap}, \mathrm{mov}\}$. The normalised observations are

$$\widetilde{r}^{(b)}_{G,s,h,k,t} = \frac{r^{(b)}_{G,s,h,k,t}}{\max\{a_{s,b}, \varepsilon\}}, \tag{27}$$

where $\varepsilon > 0$ is a numerical floor. The scales are fixed for subsequent DEV and TEST computations.

B.2 Empirical Wasserstein Distance

For each game, solver, iteration, and source signal, we form the empirical response measure

$$\mu^{(b)}_{G,s,t} = \frac{1}{N_s} \sum_{h \in H_s} \sum_{k=1}^{K} \delta_{\widetilde{r}^{(b)}_{G,s,h,k,t}}, \qquad N_s = |H_s| K. \tag{28}$$

Each configuration–initialisation pair contributes equally within its solver family.

For two games G and G', the one-dimensional Wasserstein-1 distance between the corresponding empirical measures is

$$W_1\left(\mu^{(b)}_{G,s,t}, \mu^{(b)}_{G',s,t}\right) = \frac{1}{N_s}\sum_{j=1}^{N_s}\left|\widetilde{r}^{(b)\uparrow}_{G,s,t,j} - \widetilde{r}^{(b)\uparrow}_{G',s,t,j}\right|, \tag{29}$$

where $\widetilde{r}^{(b)\uparrow}_{G,s,t,j}$ denotes the j-th ordered observation.

The behavioural distance aggregates these discrepancies across solver families, iterations, and source signals:

$$d_{\mathrm{beh}}(G, G') = \frac{1}{|\mathcal{S}|T|\mathcal{B}_{\mathrm{src}}|}\sum_{s\in\mathcal{S}}\sum_{t=1}^{T}\sum_{b\in\mathcal{B}_{\mathrm{src}}} W_1\left(\mu^{(b)}_{G,s,t}, \mu^{(b)}_{G',s,t}\right). \tag{30}$$

The aggregation assigns equal aggregate weight to each solver family and source signal, independently of the number of configurations within a family.

The empirical measures preserve time-indexed marginal distributions of exploitability and movement. Configuration identity, initialisation identity, joint associations between the two signals, and cross-time trajectory correspondence are excluded from the distance definition. Higher-order trajectory properties, including recurrence and reversal, are consequently available as withheld behavioural quantities for downstream evaluation.

B.3 Oracle Atlas Construction

Let $\mathcal{L} = \{G_i\}_{i=1}^{L}$ denote the fixed TRAIN landmark set, with pairwise behavioural-distance matrix

$$D_{ij} = d_{\mathrm{beh}}(G_i, G_j). \tag{31}$$

Classical multidimensional scaling (MDS) constructs the centred kernel

$$J = I_L - \frac{1}{L}\mathbf{1}_L\mathbf{1}_L^\top, \qquad B_{\mathcal{L}} = -\frac{1}{2}J(D\odot D)J, \tag{32}$$

where $\odot$ denotes element-wise multiplication.

Let U_d contain the eigenvectors corresponding to the d largest retained positive eigenvalues, and let Λ_d be the associated diagonal eigenvalue matrix. The landmark coordinates are

$$Z^\star_{\mathcal{L}} = U_d\Lambda_d^{1/2}. \tag{33}$$

For a game G outside the landmark set, define

$$\begin{aligned} v_i(G) &= d_{\mathrm{beh}}(G, G_i)^2, \quad \bar{v}(G) = \frac{1}{L}\sum_{i=1}^{L} v_i(G), \\ m_i &= \frac{1}{L}\sum_{j=1}^{L} D_{ij}^2, \qquad\qquad \bar{m} = \frac{1}{L^2}\sum_{i,j} D_{ij}^2. \end{aligned} \tag{34}$$

Its centred similarities to the landmarks are

$$b_i(G) = -\frac{1}{2}\left(v_i(G) - \bar{v}(G) - m_i + \bar{m}\right). \tag{35}$$

Writing $\mathbf{b}(G) = (b_1(G), \ldots, b_L(G))^\top$, the out-of-sample coordinates are obtained as

$$z^\star(G) = \Lambda_d^{-1/2}U_d^\top\mathbf{b}(G) \in \mathbb{R}^d. \tag{36}$$

The response normalisation, landmark set, and MDS eigensystem are estimated exclusively from TRAIN. The resulting coordinate system is fixed for recogniser supervision and downstream evaluation.

Table 4: **Geometry preservation on DEV.** Pairwise-distance Spearman correlation (higher is better). Learned-atlas results are reported as mean ± standard deviation across three seeds.

Comparison	Spearman ↑
d_{beh} vs. $z^\star$	0.991
$z^\star$ vs. $\hat{z}$	0.920 ± 0.012
d_{beh} vs. $\hat{z}$	0.917 ± 0.012

Table 5: **Oracle atlas dimensionality ablation on DEV.** Distance Spearman measures agreement between the raw behavioural distance and the oracle Euclidean distance over DEV–DEV game pairs. Strict LOSO results report macro-average NMAE across five behavioural targets and seven excluded solvers using a 10-nearest-neighbour median readout. Results are mean ± sample standard deviation over three paired label seeds. The raw behavioural distance provides an uncompressed reference.

	Geometric fidelity	Strict LOSO NMAE ↓		
Representation	Spearman ↑	20 labels	102 labels	205 labels
Raw distance	1.0000	0.8683 ± 0.0731	0.6888 ± 0.0357	0.6264 ± 0.0674
Oracle 4D	0.9806	0.8617 ± 0.0656	0.6896 ± 0.0314	0.6247 ± 0.0682
Oracle 8D	0.9877	0.8634 ± 0.0710	0.6866 ± 0.0328	0.6219 ± 0.0678
Oracle 12D (adopted)	0.9906	0.8655 ± 0.0714	0.6869 ± 0.0333	0.6243 ± 0.0671
Oracle 24D	0.9939	0.8700 ± 0.0712	0.6894 ± 0.0349	0.6285 ± 0.0620
Oracle 32D	0.9942	0.8712 ± 0.0714	0.6916 ± 0.0352	0.6297 ± 0.0604

B.4 Geometry Preservation

We assess the extent to which the two-stage construction preserves the geometry induced by the original behavioural distance. Table 4 reports pairwise-distance Spearman correlations on DEV, measuring distance-order agreement across the raw behavioural distance d_{beh}, the oracle atlas $z^\star$, and the payoff-derived atlas $\hat{z}$.

The oracle atlas retains the distance ordering of the original behavioural geometry with a Spearman correlation of 0.993, indicating limited distortion under 12-dimensional compression. The payoff-derived representation achieves a correlation of 0.920 ± 0.012 with the oracle atlas and 0.917 ± 0.012 with the raw behavioural distance. These results demonstrate that the recogniser recovers the broad relational structure of the behavioural geometry from game payoffs, providing a geometric basis for the downstream behavioural generalisation and cartographic analyses.

B.5 Effect of Atlas Dimensionality

We examine the effect of oracle embedding dimensionality under the Current-fullH protocol. We consider $d \in \{4, 8, 12, 24, 32\}$, using the same TRAIN landmarks and successive prefixes of the retained MDS eigensystem. This ablation evaluates the compression of the empirical behavioural distance into oracle coordinates $z^\star$; it does not involve training separate payoff recognisers at different dimensions.

We assess geometric fidelity through the Spearman correlation between the original behavioural distances and Euclidean distances in the oracle space, evaluated over DEV–DEV game pairs. Predictive utility is assessed through strict leave-one-solver-out (LOSO) transfer. For each excluded solver, the source atlas is constructed from the remaining six solvers, and a $k = 10$ nearest-neighbour median readout predicts its five panel-mean behavioural targets: gap AUC, late gap, late movement, recurrence, and reversal. All dimensions share the same labelled subsets and evaluation protocol. We report macro-average normalised mean absolute error (NMAE), averaging equally across the seven excluded solvers within each label seed, followed by the mean and sample standard deviation across three seeds.

Table 5 shows that the behavioural distance admits substantial low-dimensional compression. Even the four-dimensional oracle preserves the pairwise distance ordering with a Spearman correlation

of 0.9806, increasing to 0.9906 at 12 dimensions and 0.9942 at 32 dimensions. The corresponding strict-LOSO results remain closely comparable across the evaluated dimensions and label budgets. In particular, the eight-dimensional oracle attains slightly lower aggregate prediction errors than the adopted 12-dimensional representation, although the observed differences are small and no statistical significance test is performed.

We retain the 12-dimensional atlas as a compromise between compactness and geometric fidelity. It preserves more of the empirical distance ordering than the smaller representations, while additional dimensions yield only modest geometric improvements without corresponding gains in aggregate solver transfer. The comparable predictive performance of the four- and eight-dimensional representations further indicates that the information supporting these downstream readouts is retained under substantial compression.

This comparison isolates oracle dimensionality using fixed TRAIN-derived constructions. The effect of dimensionality on payoff-only coordinate recovery is not evaluated here. The distance correlation reported in Table 5 uses DEV–DEV pairs; the DEV–TRAIN-landmark correlation reported elsewhere therefore follows a different reference-pair protocol.

C Additional Experimental Results

This appendix provides supplementary evidence for behavioural and solver generalisation, numerical strategic invariance, and cartographic evaluation.

C.1 Behavioural and Solver Generalisation

The transfer experiments evaluate frozen representations using target-label budgets of 20, 102, and 205 games. Behaviour transfer includes trajectory properties withheld from distance construction, while strict leave-one-solver-out (LOSO) evaluation excludes the target solver from the source distance, oracle embedding, and recogniser supervision. The target solver's responses are introduced only through the downstream labelled subset.

All methods share the same target-label subsets. Frozen representations are evaluated using a ten-nearest-neighbour median readout and Ridge regression, probing local behavioural similarity and linear accessibility, respectively. Performance is measured using the normalised mean absolute error defined in Section 4. Reported uncertainty reflects the variation across the three evaluation seeds.

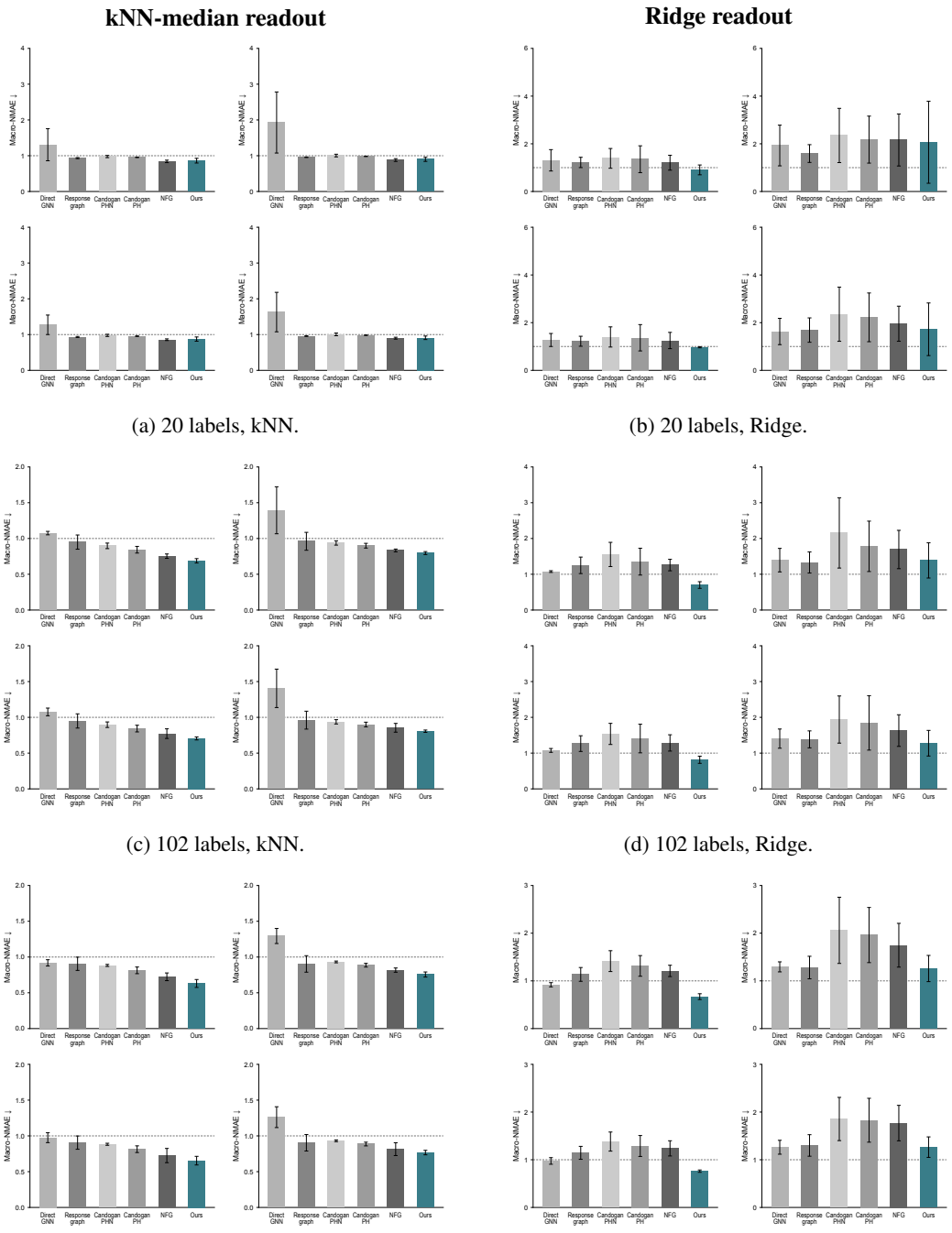
Nearest-neighbour transfer. Figure 6 provides the complete three-budget comparison. The evaluation covers source and withheld behavioural quantities under both full-library and strict LOSO settings.

Under strict LOSO evaluation, the learned atlas obtains overall NMAE 0.889, 0.745, and 0.699 at the three respective budgets. At 205 labels, the corresponding error is approximately 30% below the constant reference. The source-only raw behavioural distance achieves macro-average NMAE 0.626 at the same budget, providing a reference for the information retained before embedding and payoff-only recovery.

Linear accessibility. Figure 6 reports the corresponding Ridge comparison. For source behavioural quantities, the learned atlas obtains NMAE 0.906, 0.699, and 0.670 at 20, 102, and 205 labels, respectively. The linear readout evaluates whether behavioural information remains accessible through a simple mapping from the frozen representation.

C.2 Effect of Behavioural Pretraining

We examine whether the transfer performance of the learned behavioural atlas can be explained by source-behaviour pretraining alone. To this end, we compare the learned 12-dimensional atlas with Ordinary128-20k, a payoff encoder pretrained to regress summaries of the same permissible source gap and movement responses. Ordinary128 uses the downstream performence targets, whereas the atlas recogniser is supervised by coordinates obtained from the Current-fullH behavioural distance. The two approaches therefore differ in how the source behavioural information is organised into a representation.



(a) 20 labels, kNN.

(b) 20 labels, Ridge.

(c) 102 labels, kNN.

(d) 102 labels, Ridge.

(e) 205 labels, kNN.

(f) 205 labels, Ridge.

Figure 6: **Behavioural and solver generalisation across readouts and label budgets.** Comparison of kNN-median and Ridge readouts using 20, 102, and 205 target-labelled games. Bars report macro-average NMAE (lower is better), with error bars indicating standard deviations across three seeds. All methods are evaluated using matched target-label subsets.

Table 6: **Effect of behavioural pretraining on withheld-behaviour transfer.** Full-solver prediction of panel-mean recurrence and reversal on DEV. All methods use a 10-nearest-neighbour median readout with paired target-label subsets. Values are macro-average NMAE (lower is better), reported as mean ± sample standard deviation over three label seeds. Ordinary128-20k uses behavioural-summary pretraining, while the learned atlas uses Current-fullH oracle-coordinate supervision.

Representation	20 labels	102 labels	205 labels
Oracle atlas $z^\star$	0.8821 ± 0.0707	0.7380 ± 0.0326	0.6747 ± 0.0456
Ordinary128-20k	0.9494 ± 0.0175	0.8568 ± 0.0047	0.8520 ± 0.0040
Learned atlas $\hat{z}$	$\mathbf{0.8893 \pm 0.0573}$	$\mathbf{0.7464 \pm 0.0206}$	$\mathbf{0.7058 \pm 0.0354}$

Table 7: **Behavioural pretraining under strict LOSO evaluation on DEV.** Each fold excludes the target solver from representation construction and predicts its five panel-mean behavioural quantities. Entries report mean ± sample standard deviation across three seeds, after equally averaging the seven solver folds within each seed. All methods use identical label subsets and a 10-nearest-neighbour median readout. Ordinary and learned-atlas encoders use matched source-training games, backbone width and depth, and optimization budgets. Lower NMAE is better.

Representation	20 labels	102 labels	205 labels
Measured-behaviour references			
Raw behavioural distance	0.8683 ± 0.0731	0.6888 ± 0.0357	0.6264 ± 0.0674
Oracle atlas $z^\star$	0.8655 ± 0.0714	0.6869 ± 0.0333	0.6243 ± 0.0671
Payoff-only representations			
Ordinary128-20k	0.9352 ± 0.0119	0.8158 ± 0.0134	0.8026 ± 0.0154
Learned atlas $\hat{z}$	$\mathbf{0.8885} \pm 0.0550$	$\mathbf{0.7455} \pm 0.0190$	$\mathbf{0.6991} \pm 0.0487$

Matched pretraining and evaluation. Both models use 20,000 TRAIN games, payoff encoders with width 128 and depth 2, and matched optimisation budgets. The atlas recogniser outputs 12-dimensional coordinates, while Ordinary128 provides frozen 128-dimensional features. Downstream evaluation uses Euclidean nearest-neighbour retrieval with $k = 10$ and a median target readout. Both methods share the same labelled subsets at budgets of 20, 102, and 205 games, together with three paired label seeds. Their representations remain frozen throughout downstream evaluation.

For strict leave-one-solver-out (LOSO) evaluation, the excluded solver is removed from source supervision, normalisation, and model selection for both methods. The learned atlas is constructed from the remaining six solvers, including its behavioural distance, oracle embedding, and payoff recogniser. Ordinary128 likewise uses only the permissible source responses from these six solvers. Both representations are evaluated on the five panel-mean behavioural targets of the excluded solver. This protocol controls for the availability of source behavioural supervision while retaining the different pretraining objectives of the two representations.

Transfer to withheld behavioural quantities. We first evaluate recurrence and reversal, which are excluded from the source distance and are computed from the full configuration–initialisation panel. Table 6 reports the corresponding full-solver results. The learned atlas achieves lower macro-average NMAE than Ordinary128 at all three label budgets. The oracle atlas provides a reference for the additional approximation introduced by payoff-only coordinate recovery.

Figures 7 and 8 provide a complementary visual and predictive comparison between structural taxonomy and behavioural cartography for the two withheld trajectory properties.

Transfer to excluded solver mechanisms. We next evaluate strict LOSO transfer across all seven solver primitives. Each fold predicts gap AUC, late gap, late movement, recurrence, and reversal for the solver excluded from representation construction. Macro-average NMAE is first computed over the valid behavioural targets within each solver and then averaged equally across the seven solver folds. Table 7 reports the results at all three label budgets.

The learned atlas achieves lower aggregate prediction error than Ordinary128 across all three budgets. At 205 labels, its NMAE of 0.6991 represents a 12.9% reduction relative to Ordinary128. The raw

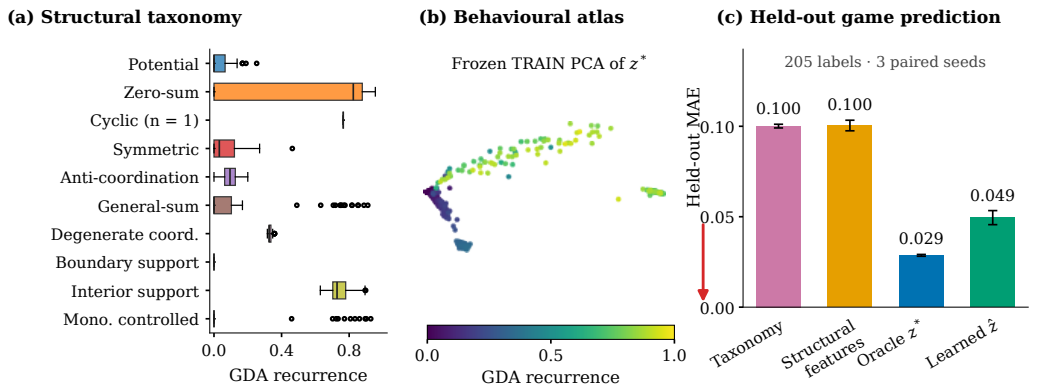


Figure 7: **Structural taxonomy and behavioural cartography of held-out recurrence.** (a) Distribution of measured GDA recurrence within each structural game family, revealing substantial within-family behavioural variation. (b) The same games projected onto the frozen TRAIN-fitted PCA of the oracle behavioural atlas and coloured by measured recurrence. Recurrence is excluded from atlas construction, and the two-dimensional projection is used solely for visualisation. (c) Held-out recurrence prediction using taxonomy labels, structural features, oracle coordinates $z^\star$, and learned coordinates $\hat{z}$. Bars show mean MAE over three paired label seeds at 205 labels, with error bars indicating sample standard deviation. The readouts use the full 12-dimensional coordinates. Lower is better.

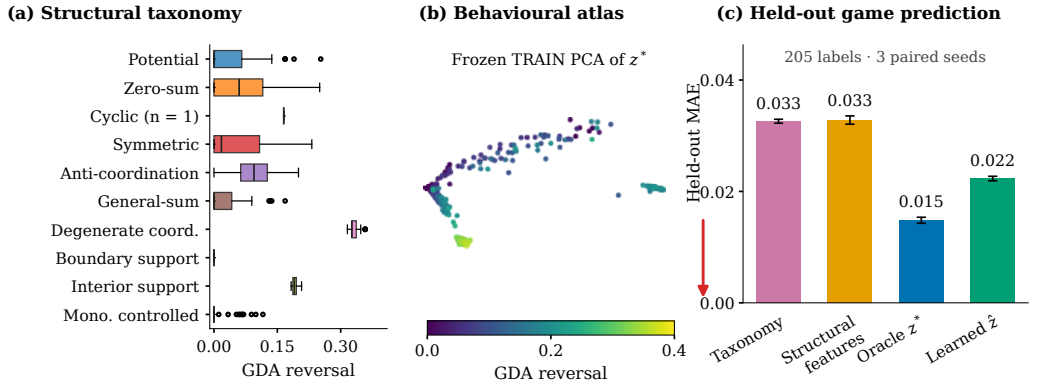


Figure 8: **Structural taxonomy and behavioural cartography of held-out reversal.** (a) Distribution of measured GDA reversal within each structural game family, showing behavioural variation both within and across structural families. (b) The same games projected onto the frozen TRAIN-fitted PCA of the oracle behavioural atlas and coloured by measured reversal. Reversal is excluded from atlas construction, and the two-dimensional projection is used solely for visualisation. (c) Held-out reversal prediction using taxonomy labels, structural features, oracle coordinates $z^\star$, and learned coordinates $\hat{z}$. Bars show mean MAE over three paired label seeds at 205 labels, with error bars indicating sample standard deviation. The readouts use the full 12-dimensional coordinates. Lower is better.

behavioural distance and oracle atlas attain NMAE values of 0.6264 and 0.6243, respectively, indicating that much of the additional error arises during payoff-only recovery. The small difference between the raw distance and oracle results is consistent with the dimensionality analysis in Appendix B.5.

Table 8 further reports the 205-label results for each excluded solver. The learned atlas has lower seed-averaged NMAE than Ordinary128 in all seven folds, although the magnitude of the difference

Table 8: **Per-solver strict LOSO transfer at 205 labels.** Each entry is the mean NMAE over three paired label seeds, averaged over the five panel-mean behavioural targets of the excluded solver. Lower is better. The final row gives the equal-weight macro-average across the seven folds.

Excluded solver	Ordinary128-20k	Learned atlas $\hat{z}$
GDA	0.6825	**0.5377**
Entropic mirror descent	0.9965	**0.9313**
Damped proximal	0.9949	**0.8809**
Averaging/anchoring	0.6366	**0.4583**
Extragradient	0.8331	**0.7418**
Optimistic gradient	0.5690	**0.4879**
Fictitious play	0.9056	**0.8559**
Macro-average	0.8026	**0.6991**

varies across solver mechanisms. The comparison therefore supports the aggregate transfer result while revealing heterogeneous solver-specific benefits.

Interpretation and limitations. The comparison provides evidence that atlas-coordinate supervision offers additional predictive utility over the evaluated behavioural-summary pretraining baseline. Both models receive the same permissible source behavioural information and are evaluated using matched downstream protocols. Their different objectives induce different organisations of that information, with the atlas explicitly approximating a distance-derived coordinate system shared across solver mechanisms.

C.3 Strategic Invariance

We evaluate invariance of the frozen behavioural atlas under strategically equivalent presentations of the same game. For each 3×3 game, the transformation group consists of 72 presentations induced by independent action permutations and optional player exchange. These transformations preserve the underlying strategic interaction while changing its payoff representation.

Recogniser and downstream predictions. We evaluate the complete transformation orbit for each of the 5,371 games in the established TEST split across three recogniser seeds. For each game, coordinate discrepancy is measured by the maximum Euclidean distance between its original representation and the representations of its transformed presentations.

The mean per-game maximum discrepancy is 4.660×10^{-6}, with a global maximum of 6.911×10^{-5} under float32 inference. Additional float64 diagnostics reduce the maximum discrepancies of the three worst-case games to at most 5.044×10^{-14}, consistent with floating-point rounding effects.

We further examine whether these numerical discrepancies affect frozen downstream predictions. Using the 205-label budget and ten-neighbour median readout, predictions remain identical across all transformations for withheld behavioural quantities across three seeds and for a representative strict LOSO experiment excluding extragradient. The selected neighbour sets also remain unchanged.

Solver-response protocol. We separately assess numerical invariance of the empirical behavioural distance, including solver execution, initialisation, and response aggregation. This experiment uses 16 TEST games, all 31 solver configurations, 19 initialisations, and 200 optimisation steps. Six fixed transformations are applied to each game.

Both transformed and fixed initialisation panels are evaluated over 4,320 transformed game-pair comparisons. The protocols yield identical aggregate errors, with mean absolute discrepancy 1.482×10^{-3} and maximum discrepancy 2.305×10^{-2}.

The largest discrepancies arise from large-step extragradient dynamics, where perturbations initially at floating-point precision are amplified over successive updates. The original fictitious-play implementation also employs index-dependent tie-breaking, which can affect equivariance when best responses are tied.

Table 9: **Strategic invariance of the frozen behavioural atlas.** Coordinate errors are measured over transformation orbits, while behavioural-distance errors are absolute discrepancies over transformed game pairs.

Evaluation	Mean error	Maximum error
Recogniser coordinates (three seeds)	4.660×10^{-6}	6.911×10^{-5}
Withheld behaviour predictions (three seeds)	0	0
Withheld solver predictions (extragradient)	0	0
Behavioural distance, matched initialisations	1.482×10^{-3}	2.305×10^{-2}
Behavioural distance, fixed initialisation panel	1.482×10^{-3}	2.305×10^{-2}

Table 9 summarises the numerical audit.

These results establish numerical invariance of the payoff recogniser to float32 precision and identical downstream predictions within the evaluated transformation orbits. The separate solver-response audit identifies finite-precision discrepancies in the empirical behavioural distance.

C.4 Controlled Structural Paths

Controlled structural paths. We further examine how controlled changes in game structure are reflected in the atlas and its associated behavioural responses. We consider two withheld dynamical quantities: recurrence, which measures repeated visits to previously encountered strategy states, and reversal, which captures changes in the direction of successive strategy updates. Both are evaluated from solver trajectories and excluded from the behavioural distance used to construct the atlas.

Figure 5 follows a representative mixture path

$$G(\lambda) = (1 - \lambda)G_A + \lambda G_B, \qquad \lambda \in [0, 1],$$

which continuously interpolates between two games while changing their strategic structure. Successive games trace an ordered trajectory through the behavioural atlas, accompanied by systematic changes in measured recurrence and reversal. Oracle and learned atlas readouts recover the overall progression of both withheld behaviours, including their crossings of thresholds fixed on TRAIN. This shows that the atlas captures behavioural variation induced by continuous structural changes in the underlying game, rather than only separating discrete game classes.

All quantitative distances are computed in the full-dimensional representation. The two-dimensional projection is used exclusively for visualisation.

Across the evaluated paths, the oracle atlas preserves local behavioural-distance ordering closely, achieving Spearman correlations between 0.974 and 0.977 for the representative paths reported in Figure 5. The learned atlas achieves correlations between 0.675 and 0.743 on the same paths.

Under the full solver-configuration panel, the overall recurrence-prediction NMAE is 0.428, 0.386, and 0.568 for the raw behavioural distance, oracle atlas, and learned atlas, respectively. These results indicate that the oracle coordinates retain substantial information about local behavioural changes, with additional distortion introduced by payoff-only recognition. The representative paths in Figure 5 illustrate both recovered behavioural variation and residual discrepancies in transition magnitude.

C.5 Payoff-Scale Sensitivity

We additionally examine finite-time sensitivity to payoff magnitude under fixed solver configurations. Starting from a normalised base game G_0, we construct $G(c) = cG_0$ without subsequent per-game normalisation. The solver step-size panel remains fixed, so scaling the payoffs changes the effective update magnitude while preserving the relative payoff structure. Figure 9 illustrates the resulting variation in measured dynamics and atlas-based behavioural readouts. This diagnostic is evaluated separately from the normalised structural-path protocol.

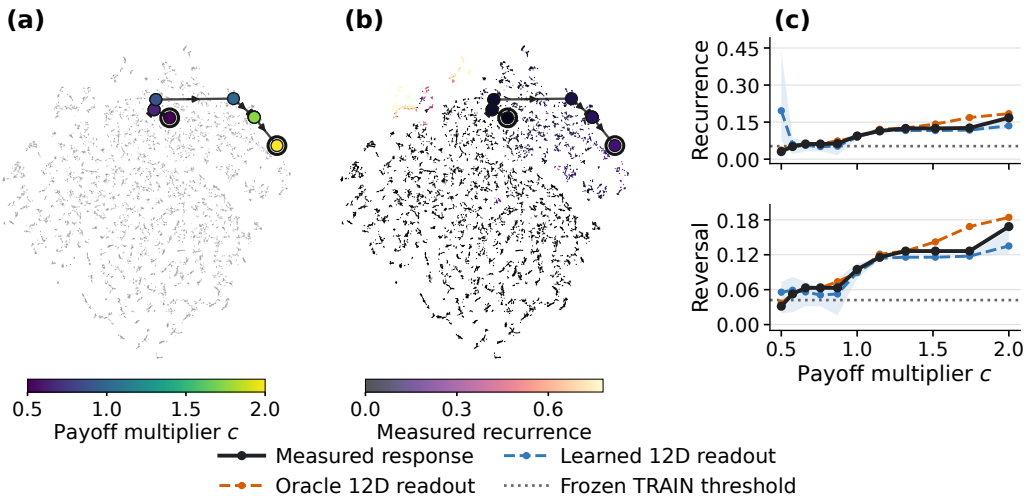


Figure 9: **Payoff-scale sensitivity under fixed solver configurations.** A normalised base game's payoffs are multiplied by c without subsequent per-game normalisation. The panels show the resulting atlas trajectory, measured recurrence, and oracle and learned behavioural readouts.

C.6 Cross-Size Generalisation and Transfer

Protocol. We train the behavioural encoder on 20,000 3×3 games, using 512 source-domain validation games for checkpoint selection. We then apply the frozen encoder to 4×4 and 5×5 games without updating its parameters. The scaling and 12-dimensional MDS coordinate system are also fitted using source training games only and remain fixed. Thus, the embedding z of a new game is inferred directly from its payoff matrices; neither the encoder nor the coordinate system requires retraining for the new size. All sizes use the same solver step-size settings, rollout horizon $T = 200$, and initialization budget $K = 19$.

We distinguish two transfer settings. In *source-label zero-shot transfer*, predictions use 2,048 labelled 3×3 reference games and require no target-size labels. In *same-size reference adaptation*, predictions instead use 2,048 labelled reference games of the target size, while the encoder remains frozen. Both settings use a $k = 10$ nearest-neighbour median readout and 512 disjoint query games per target size. The latter requires target-domain behavioural labels, although no neural-network retraining is performed, and should therefore not be interpreted as label-free transfer.

Comparison with baselines. Table 10 reports NMAE averaged over 35 solver–behaviour targets. For neural methods, each entry is the mean and sample standard deviation of three independently trained seeds; predictions are not ensembled. The standard deviations describe variation across training seeds, not uncertainty from sampling evaluation games. Response graph and Candogan are deterministic evaluations and consequently have no training-seed standard deviation in this table.

With source 3×3 reference labels, our method has the lowest mean NMAE among the evaluated payoff-only methods at each tested size. With same-size reference labels, it has the lowest reported mean at 3×3 and 4×4, while response graph has a lower reported value at 5×5 (0.5892 versus 0.6443 ± 0.0643). The 5×5 result varies across our training seeds; one of our three seeds scores below the response-graph value. These are descriptive comparisons, and this experiment does not establish a statistically significant ranking between methods.

Our encoder achieves lower average zero-shot NMAE than the evaluated NFG and response-graph baselines at every tested size. However, size compatibility is not unique to our method: both baselines can also be evaluated across sizes. With same-size reference labels, our method leads at 3×3 and 4×4, whereas response graph performs better at 5×5. These comparisons are descriptive and do not establish statistical significance.

Table 10: Cross-size behaviour prediction NMAE. Neural entries report the mean and sample standard deviation across three training seeds. Response graph and Candogan are deterministic evaluations and therefore have no training-seed standard deviation. Source-reference transfer uses no target-size labels; same-size adaptation uses 2,048 target-size reference labels while keeping the encoder frozen. The 3×3 column contains newly generated evaluation games.

Method	3×3	4×4	5×5
Source-reference transfer (zero-shot)			
Ours	$\mathbf{0.5810 \pm 0.0643}$	$\mathbf{0.6040 \pm 0.0489}$	$\mathbf{0.6739 \pm 0.0321}$
NFG	0.6329 ± 0.0467	0.6899 ± 0.0707	0.7600 ± 0.0622
Response graph	0.7022	0.7699	0.7898
Same-size reference adaptation			
Ours	$\mathbf{0.5201 \pm 0.0558}$	$\mathbf{0.5691 \pm 0.0576}$	0.6443 ± 0.0643
NFG	0.5693 ± 0.0276	0.6124 ± 0.0280	0.6642 ± 0.0326
Response graph	0.5653	0.6146	**0.5892**
Candogan PH	0.5832	0.6729	0.7306
Candogan PHN	0.7019	0.7597	0.7566

Transfer and increasing game size. Our encoder and coordinate system are fitted on 3×3 source games and remain frozen when applied to larger games. In source-reference transfer, no target-size behavioural labels are used. Its mean NMAE increases from 0.5810 ± 0.0643 on newly generated 3×3 games to 0.6040 ± 0.0489 on 4×4 games and 0.6739 ± 0.0321 on 5×5 games. The same-size reference setting shows a similar increase, from 0.5201 ± 0.0558 to 0.5691 ± 0.0576 and 0.6443 ± 0.0643. Thus, the observed average prediction error grows with game size over the tested range, although the table alone does not establish a general scaling law.

Same-size reference adaptation uses 2,048 labelled games of each target size while leaving the encoder frozen. Its improvement over source-reference transfer demonstrates that target-size labels help the readout; it does not demonstrate few-shot transfer or eliminate the cost of obtaining those labels. Moreover, the new 3×3 evaluation games already differ from the source training distribution, and the fixed $K = 19$ initialization budget covers a smaller fraction of larger games' possible initializations. The observed decline therefore cannot be attributed solely to matrix size.

Degradation with increasing size difference. Performance deteriorates as the evaluation size increases: zero-shot NMAE rises from 0.5810 to 0.6040 and 0.6739, and the same trend appears with same-size reference labels. For our embedding, same-size NN10 overlap also decreases from 0.1107 to 0.0384 and 0.0277. Thus, transfer remains useful at 5×5, but neighbourhood preservation and downstream prediction weaken with increasing size difference in the tested range.

This trend is descriptive rather than a causal isolation of game size. The new evaluation generator already differs from the source training distribution, and a fixed $K = 19$ covers a smaller fraction of initialization configurations at larger sizes. Moreover, three tested sizes do not establish a general scaling law. Future controls should vary target-label budgets, initialization coverage, and training-size diversity while retaining held-out larger games for evaluation.