# From Risk Scoring to Risk Allocation: A Density-Driven Framework for Diverse Monitoring in Multi-Agent Systems

**Zhaohui Wang** [1]

## Abstract

Risk monitoring in multi-agent systems is commonly built on a per-state primitive that scores each state independently and selects the top $K$. Under crowding, where many agents share the same fragility, this approach picks redundant alerts whose risks are jointly correlated, a pattern we describe as "herding in monitoring." We propose a paradigm shift from risk scoring to risk allocation, supported by two contributions. First, we identify the *Crowding Paradox*, namely that $P(\text{risk} \mid x) \propto p(x)$ rather than $1/p(x)$, so density rather than anomaly score is the operative risk signal; on financial data, density-based scoring reaches AUROC $\geq 0.94$ at 5d/10d/20d crash horizons, while five anomaly baselines all fall below $0.80$. Second, given a density-derived fragility score, we recast monitoring as combinatorial subset selection over interdependent states and map it to a QUBO objective with a $\lambda$-controlled risk–diversity tradeoff. The resulting Pareto frontier contains standard diverse-subset methods (MMR, $k$-DPP) as fixed operating points; the gain over greedy grows monotonically with scale, from $+24\%$ at $n$=15 to $+66\%$ at $n$=200; a learned $\lambda$ policy reaches $99.5\%$ of an oracle grid-search objective; and the formulation transfers to traffic and multi-agent reinforcement learning. The same QUBO instances execute without modification on Rigetti superconducting QPUs (Ankaa-3 and Cepheus-1-108Q via Amazon Braket), which we report as a compatibility property of the formulation rather than a claim of quantum advantage at this scale.

[1]University of Southern California, Los Angeles, CA, USA. Correspondence to: Zhaohui Wang <zwang000@usc.edu>.

*Proceedings of the $43^{rd}$ International Conference on Machine Learning*, Seoul, South Korea. PMLR 306, 2026. Copyright 2026 by the author(s).

## 1. Introduction

A central feature of risk monitoring in multi-agent systems is its anomaly-centric primitive: states are scored independently, and the lowest-probability states are flagged for attention. This primitive is at odds with how systemic risk is generated. In multi-agent systems such as financial markets, autonomous fleets, and cooperative robotics, risk concentrates not in rare states but in crowded ones, where many agents adopt similar strategies and share the same failure modes. Banerjee's herding model (Banerjee, 1992) shows that rational agents may ignore private information and follow their predecessors, creating information cascades. The 2007 quant meltdown provides direct evidence: quantitative funds with similar factor exposures experienced synchronized liquidations (Khandani & Lo, 2011), and subsequent studies confirm that crowding measures predict tail risk across asset classes (Brown et al., 2022; Kang et al., 2021), with fire-sale contagion amplifying losses when crowded positions unwind simultaneously (Cont & Wagalath, 2016; Greenwood et al., 2015). These findings point to what we call the **Crowding Paradox**: in contrast to anomaly-detection methods that equate risk with low-probability states (Chandola et al., 2009), systemic risk arises from *high-density* states where agents cluster, so that $P(\text{risk} \mid x) \propto p(x)$ rather than $1/p(x)$. We validate this empirically in App. A: on the 2024–2026 financial test set, GMM density and the derived Fragility Score reach AUROC values of $0.94$, $0.94$, and $0.96$ at 5d, 10d, and 20d crash horizons respectively, while five anomaly-detection baselines (Isolation Forest, OCSVM, LOF, PCA-reconstruction, and AE-reconstruction) all score below $0.80$ AUROC and below $0.06$ AUPRC.

Identifying which states are fragile is only half of the problem; the remaining half is allocating finite monitoring capacity across them. This motivates the central reframing of the paper, namely a shift from risk scoring to risk allocation. Score-and-threshold approaches implicitly assume that the $K$ riskiest states are statistically independent, an assumption that fails under crowding because crowded states share failure modes. The resulting top-$K$ selection clusters in feature space and is vulnerable to a single shock, an outcome that we describe as "herding in monitoring" since it mirrors the herding that creates fragility in the first place.

Subset selection with a diversity constraint is a *discrete optimization* problem. Rather than relying on ad-hoc post-processing such as diversity re-ranking, we formulate the problem directly as QUBO (Glover et al., 2022):

$$\min_{\mathbf{z}\in\{0,1\}^n} \underbrace{-\sum_i \mathrm{FS}(x_i)\, z_i}_{\text{risk coverage}} + \lambda \underbrace{\sum_{i<j} S_{ij}\, z_i z_j}_{\text{diversity}} + P\Big(\sum_i z_i - K\Big)^2 \quad (1)$$

where $\mathrm{FS}(x_i)$ is the Fragility Score of state $x_i$, $S_{ij}$ measures pairwise similarity, and $P$ enforces cardinality $K$.

This formulation offers three advantages over score-and-threshold approaches: (i) it jointly optimizes risk and diversity in a single objective; (ii) it is agnostic to the solver backend, so that classical heuristics, exact methods, and quantum processors can all solve the same QUBO; (iii) it provides a principled $\lambda$-parameterized tradeoff between risk concentration and coverage diversity.

**Contributions.** **(1) Conceptual.** We identify and empirically validate the *Crowding Paradox*, namely that in multi-agent systems risk concentrates in high-density states rather than anomalies, so that density rather than anomaly score is the operative risk signal (App. A). **(2) Reframing.** We recast risk-aware monitoring as a combinatorial allocation problem over interdependent states, rather than as independent per-state scoring (§2). **(3) Formulation.** We show that this allocation maps to QUBO, yielding a single quadratic objective that any QUBO-compatible solver, including classical heuristics, exact methods, quantum annealers, and gate-model algorithms, can attack without reformulation (§2, §3). **(4) Empirical.** The formulation produces a controllable risk–diversity Pareto frontier on which standard methods (greedy, MMR, $k$-DPP) sit as fixed points; the gain over greedy grows with scale, from $+24\%$ at $n$=15 to $+66\%$ at $n$=200; the formulation generalizes across finance, traffic, and multi-agent reinforcement learning; and the same QUBO instances are executable on real Rigetti QPUs (§4, App. D–H).

The substantive contribution of this work is the reframing of risk monitoring rather than the QUBO encoding itself; the QUBO formulation is used because it is the smallest quadratic encoding that supports the reframing without auxiliary variables or continuous relaxation.

## 2. Problem Formulation

**Setting.** Consider a multi-agent system producing a time series of market states $\{x_t\}_{t=1}^T$, where $x_t \in \mathbb{R}^d$ encodes features such as returns, volatilities, and cross-asset correlations. A density model (GMM with $k$=5 components) trained on historical data provides $p(x_t)$, and conditional variance $\mathrm{Var}(\Delta x|x_t)$ is estimated from a rolling 20-day window. High-density states with suppressed variance are most vulnerable: agents in crowded positions reduce activity, compressing variance while fragility accumulates (Brown et al., 2022). We capture this mechanism with the **Fragility Score**:

$$\mathrm{FS}(x) = \frac{p(x)}{\mathrm{Var}(\Delta x|x)} \quad (2)$$

States with high density but low variance, the "calm before the storm," receive the highest fragility scores. For numerical stability, we rank-normalize FS to $[0, 1]$ across the full test set.

**Candidate pool.** From the test set ($T$=523 trading days), we select the top-$n$ states by FS as the candidate pool. We use $n$=15 as the primary setting to enable exact enumeration as ground truth, allowing controlled comparison across solvers under identical QUBO instances and isolating solver behavior from approximation error. We also evaluate scaling to $n$=20, 25, and 30 (§D). The monitoring task is to choose $K$=8 states that balance high aggregate fragility with diversity.

**QUBO construction.** The QUBO formulation is not an arbitrary encoding but the smallest quadratic form that simultaneously expresses (i) the additive risk-coverage reward $\sum_i \mathrm{FS}(x_i)z_i$, (ii) the pairwise diversity penalty $\sum_{i<j} S_{ij}z_iz_j$, and (iii) the exact cardinality constraint $\sum_i z_i = K$ as a polynomial, with no higher-order term, no auxiliary variable, and no continuous relaxation. We normalize FS within the candidate pool to $[0, 1]$ and compute pairwise similarity $S_{ij}$ via an RBF kernel on the $d$-dimensional feature vectors, also normalized to $[0, 1]$. The QUBO matrix $Q$ is obtained by expanding Eq. 1:

$$Q_{ii} = -\widehat{\mathrm{FS}}_i + P(1 - 2K) \quad (3)$$

$$Q_{ij} = \lambda S_{ij} + 2P \quad (i < j) \quad (4)$$

where $P$=2.0 is the cardinality penalty strength and $\lambda \in [0, 2]$ controls the risk-diversity tradeoff. The constant $P$ is chosen so that the marginal cardinality penalty $2P = 4$ exceeds $\lambda \cdot \max_{i<j} S_{ij} \leq \lambda$ for $\lambda \leq 2$, which guarantees that no swap that violates $|\mathcal{S}| = K$ can lower the objective at $\lambda \leq 0.5$; at $\lambda$=1.0 the bound is tight and we observe a small relaxation of the cardinality (mean $|\mathcal{S}| = 7.2$ in Table 1), indicating the soft penalty is approaching its operating limit. When $\lambda$=0, the problem reduces to pure fragility maximization (equivalent to greedy); as $\lambda$ increases, the solver trades fragility for diversity. This QUBO can equivalently be expressed as an Ising Hamiltonian via $z_i = (1 + s_i)/2$ (Lucas, 2014), enabling direct mapping to quantum hardware.

**Generalization of standard methods (informal).** Greedy top-$K$ selection corresponds to the special case $\lambda$=0 in Eq. 1, since the diversity term vanishes and the cardinality

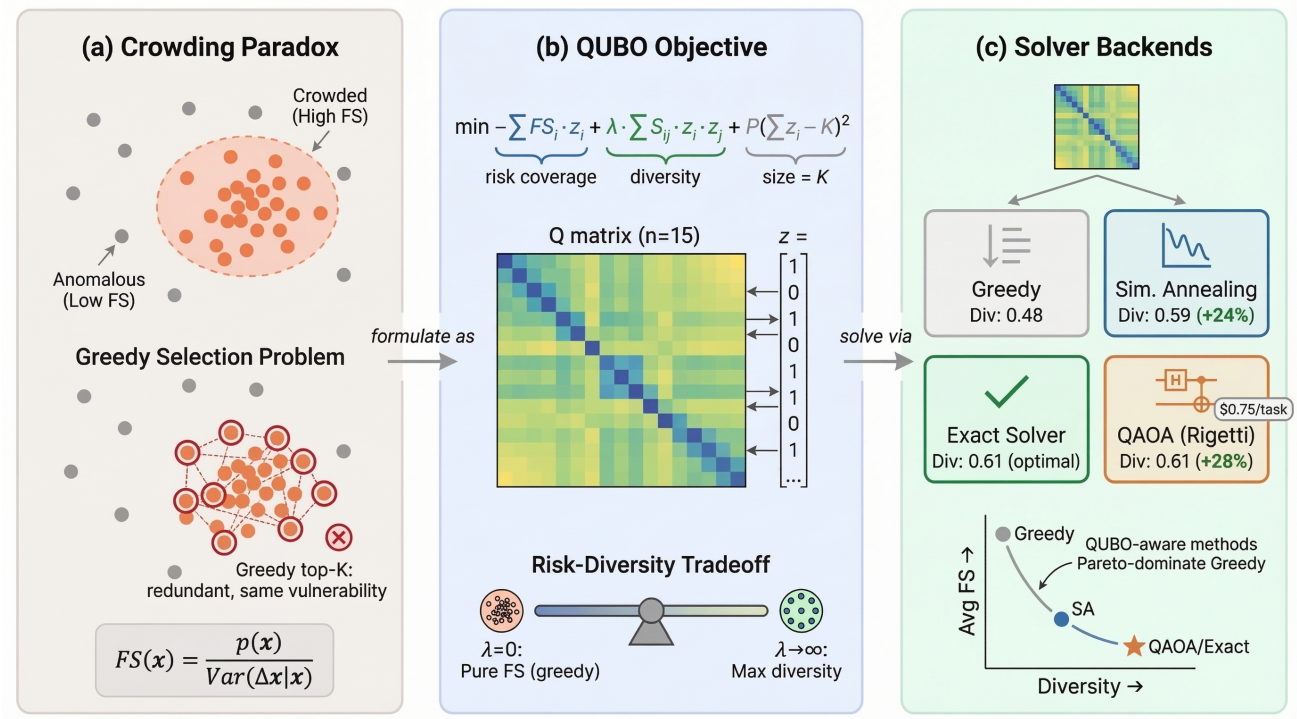


*Figure 1.* From scoring to allocation. Unlike scoring-based pipelines that threshold each state independently, the proposed formulation treats monitoring as a *joint decision* over interdependent states. (a) Multi-agent crowding creates high-density, low-variance states (high Fragility Score); greedy top-$K$ selection picks redundant states sharing the same vulnerability. (b) The QUBO objective jointly optimizes risk coverage (blue), diversity (green), and cardinality (gray) in a single quadratic form, with $\lambda$ controlling the risk-diversity tradeoff. (c) The same QUBO instance is solved by four backends: greedy (Div 0.48), simulated annealing (Div 0.59, $+24\%$), exact solver (optimal), and QAOA on Rigetti hardware. All QUBO-aware methods Pareto-dominate greedy.

penalty enforces $|\mathcal{S}| = K$. Maximal Marginal Relevance with diversity weight $\mu$ corresponds, under iterative greedy decoding, to a $\lambda$-dependent surrogate of the same objective with effective $\lambda \approx \mu/(1-\mu)$. The $k$-DPP MAP problem under the quality-similarity kernel $L_{ij} = \widehat{\text{FS}}_i \widehat{\text{FS}}_j S_{ij}$ admits a log-determinant relaxation that, locally near the QUBO minimum, is dominated by the same pairwise similarity term. Greedy, MMR, and $k$-DPP therefore correspond to specific (sometimes implicit) $\lambda$ values of a single QUBO objective, while the QUBO admits the full continuum $\lambda \in [0, 2]$, and so the formulation is a generalization of the three. Our experiments confirm that MMR($\mu$) and $k$-DPP land on the SA Pareto curve (Tables 2, 9).

## 3. Solvers

We compare four solvers on identical QUBO instances. Algorithm 1 summarizes the end-to-end pipeline.

**Greedy top-$K$** selects the $K$ highest-FS candidates, ignoring the QUBO structure entirely. This represents current practice in risk monitoring systems.

**Simulated Annealing (SA)** uses the D-Wave Ocean SDK's `SimulatedAnnealingSampler` with 200 reads per instance (Kirkpatrick et al., 1983; D-Wave Quantum Inc., 2024).

**Algorithm 1** QUBO-based fragility-aware subset selection

**Require:** Test states $\{x_t\}$, density model $p$, variance estimates, target $K$, trade-off $\lambda$
1: Compute FS($x_t$) for all test states via Eq. 2
2: Select top-$n$ by FS as candidate pool $\mathcal{C}$
3: Compute similarity matrix $S_{ij} = \text{RBF}(x_i, x_j)$ for $x_i, x_j \in \mathcal{C}$
4: Construct QUBO matrix $Q$ with parameters $\lambda, P, K$
5: Solve $\mathbf{z}^* = \arg\min_{\mathbf{z}} \mathbf{z}^\top Q\, \mathbf{z}$ via chosen solver
6: **return** Selected subset $\mathcal{S} = \{x_i \in \mathcal{C} : z_i^* = 1\}$

**Exact solver** enumerates all $2^n$ bitstrings for $n \leq 20$ via `dimod.ExactSolver`, providing the ground-truth optimal solution for verification.

**QAOA on quantum hardware** (App. B, I, J). We additionally execute the same QUBO on the local Braket statevector simulator and on Rigetti superconducting QPUs (Ankaa-3, retired during this work, and its successor Cepheus-1-108Q) via Amazon Braket, using $p$=1 QAOA (Farhi et al., 2014; Zhou et al., 2020) with 500 shots per submission. We re-

*Table 1.* Classical solver comparison at $n$=15, $K$=8, horizon=5d (mean±std over 5 data-resampling seeds; Greedy is $\lambda$-independent). QAOA on simulator and on Rigetti hardware (Ankaa-3, Cepheus-1-108Q) tracks the same Pareto frontier within hardware noise; full QAOA results in Table 5 (App. B).

| $\lambda$ | Method | Avg FS | Diversity | $\lvert\mathcal{S}\rvert$ |
|---|---|---|---|---|
| 0.0 | Greedy | .367±.002 | .476±.035 | 8.0 |
| | SA | .366±.002 | .487±.029 | 8.0 |
| | Exact | .367±.002 | .476±.035 | 8.0 |
| 0.5 | Greedy | .367±.002 | .476±.035 | 8.0 |
| | SA | .348±.025 | **.591**±.070 | 8.0 |
| | Exact | .348±.025 | **.591**±.070 | 8.0 |
| 1.0 | Greedy | .367±.002 | .476±.035 | 8.0 |
| | SA | .337±.032 | **.639**±.062 | 7.2 |
| | Exact | .337±.032 | **.639**±.062 | 7.2 |

port this as a compatibility check, not a quantum-advantage claim.

## 4. Experiments

**Data.** We use daily financial market data (2018–2025) comprising 18 features: equity returns and volume changes (SPY, QQQ), Treasury yields (2Y, 3M, 10Y), yield curve slope, investment-grade and high-yield credit spreads, realized volatility (5/10/20-day windows), gold and bond returns (TLT, GLD), and cryptocurrency returns (BTC, ETH). A GMM ($k$=5, full covariance) is trained on data through 2021; the test set comprises 523 trading days after 2023-12-31, with no temporal leakage. The top 15 states by rank-normalized Fragility Score form the candidate pool.

**Evaluation metrics.** For each selected subset $\mathcal{S}$: (i) **Avg FS**: mean Fragility Score (higher = more risk-aware); (ii) **Diversity**: $1 - \bar{S}$, where $\bar{S}$ is the mean pairwise RBF similarity within $\mathcal{S}$ (higher = more diverse); (iii) $|\mathcal{S}|$: number of selected states (target: $K$=8). We sweep $\lambda \in \{0.0, 0.5, 1.0\}$ and report results across 5 random seeds (80% candidate overlap) for stability. Detailed QAOA protocol (5 classical-optimizer seeds for the simulator; 3 independent QPU submissions per $\lambda$ on Cepheus-1-108Q; Ankaa-3 single-seed retained for the $n$=20 scaling check) is in App. B.

### 4.1. Main Results

Table 1 presents the core results.

**QUBO-aware methods improve diversity over both greedy selection and standard diverse-subset baselines.** At $\lambda$=0.5, SA and Exact increase diversity from 0.476 to 0.591, a relative gain of +24% over Greedy at a cost of only 5% in FS. At $\lambda$=1.0, diversity reaches 0.639, a gain of +34%. Beyond Greedy, we compare against standard diverse-subset methods in Table 2, namely Maximal Marginal Relevance (MMR) at three diversity weights, $k$-DPP greedy MAP under the quality-similarity kernel $L_{ij} = \widehat{\mathrm{FS}}_i \widehat{\mathrm{FS}}_j S_{ij}$, and a score-then-rerank baseline that takes the top-$2K$ by FS and applies MMR. The strongest of these, MMR at diversity weight 0.7, reaches diversity 0.562, which SA exceeds by +5% at $\lambda$=0.5 and by +14% at $\lambda$=1.0 at comparable Avg FS, with +19% over $k$-DPP. The remaining gain over MMR and $k$-DPP is smaller than the gain over Greedy, but constitutes a Pareto improvement on the evaluated instances and reduces the chance that selected states share an underlying shock.

*Table 2.* Diverse-subset baselines at $n$=15, $K$=8, horizon=5d (5 seeds; mean±std). MMR($\mu$) uses diversity weight $\mu \in [0, 1]$.

| Method | Avg FS | Diversity | $\lvert\mathcal{S}\rvert$ |
|---|---|---|---|
| MMR(0.3) | .365±.002 | .510±.052 | 8.0 |
| MMR(0.5) | .364±.002 | .518±.043 | 8.0 |
| MMR(0.7) | .349±.024 | .562±.067 | 8.0 |
| $k$-DPP (greedy MAP) | .360±.008 | .539±.057 | 8.0 |
| Score+MMR(0.5) | .364±.002 | .518±.043 | 8.0 |
| *QUBO SA,* $\lambda$=0.5 | .348±.025 | .591±.070 | 8.0 |
| *QUBO SA,* $\lambda$=1.0 | .337±.032 | .639±.062 | 7.2 |

*Table 3.* Scaling of the SA-over-Greedy diversity gain on the full financial test set ($\lambda$=0.5, $K$=8, horizon=5d). The relative gain grows monotonically with $n$; details in App. D.

| $n$ | 15 | 50 | 100 | 200 |
|---|---|---|---|---|
| Greedy Div | .476 | .495 | .505 | .514 |
| SA Div ($\lambda$=0.5) | .591 | .649 | .738 | **.851** |
| Gain over Greedy | +24% | +31% | +46% | **+66%** |
| SA runtime | 44ms | 0.6s | 4.1s | 56s |

**The gain grows with problem scale.** Extending the candidate pool to the full financial test set, SA at $\lambda$=0.5 achieves diversity 0.851 at $n$=200 versus Greedy's 0.514, an improvement of +66% (Table 3; details in App. D). The relative gain rises monotonically with $n$, indicating that the advantage is a scale-aware property of the formulation rather than a small-$n$ artifact, and SA runtime remains practical at 56 seconds for $n$=200.

**SA recovers the exact optimum, and standard methods occupy fixed points on the resulting frontier.** At $n$=15, SA produces solutions identical to those of the exact solver within seed variance, so the QUBO objective defines the Pareto frontier rather than approximating it heuristically. MMR and $k$-DPP sit on the same frontier as fixed operating points (Table 2): Greedy corresponds to $\lambda$=0, MMR(0.7) to approximately $\lambda$≈0.3, and $k$-DPP to approximately $\lambda$≈0.4. The QUBO formulation thus produces a controllable continuum that contains these existing methods rather than a single isolated operating point.

**The same QUBO instance executes on real quantum hardware without modification.** We additionally verify that the QUBO instances accepted by SA are accepted, unchanged, by gate-model quantum hardware: multi-seed QAOA on Rigetti Cepheus-1-108Q yields diversity $0.592 \pm 0.020$ at $\lambda$=0.5 and $0.616 \pm 0.007$ at $\lambda$=1.0, both tracking the classical Pareto frontier within hardware noise. We do not claim quantum advantage at this scale; we claim only computational compatibility, which is a property of the QUBO abstraction (details and cost itemization in App. B, I, J).

### 4.2. Robustness, Generalization, and Downstream Capture

The single-table headline above is supported by an extended set of analyses, all reported in detail in the appendix:

- **Risk–diversity Pareto front** (App. B): QUBO-aware methods Pareto-dominate Greedy across $\lambda \in \{0, 0.5, 1\}$; QAOA on both Rigetti devices tracks the SA/Exact frontier within hardware noise.
- **Cross-horizon consistency** (App. C): SA/Exact solutions are identical across 5d/10d/20d horizons (the QUBO objective is horizon-agnostic).
- **Scaling** (App. D): from $n$=15 to $n$=200, SA runtime grows 44ms $\rightarrow$ 56s and the diversity gain over Greedy grows monotonically from $+24\%$ to $+66\%$.
- **Sensitivity to $K$ and $\lambda$** (App. E): the SA-over-Greedy advantage holds across $K \in \{4, \ldots, 12\}$; a fine-grained $\lambda$ sweep shows a sharp diversity onset near $\lambda \approx 0.1$ and saturation beyond $\lambda \approx 1$.
- **Learned $\lambda$ policy** (App. F): an MLP policy mapping pool statistics $\rightarrow \lambda$ achieves $99.5\%$ of the oracle grid-search objective on held-out test data.
- **Cross-domain transfer** (App. G): on inD traffic and SMAC v2 MARL, SA recovers diversity that Greedy fails to deliver (Greedy diversity $\rightarrow 0$ on traffic; +7% on MARL where Greedy is already diverse).
- **Downstream stress capture** (App. H): on forward-looking SPY drawdown, realized volatility, temporal coverage, and labeled structural-break recall, SA at $\lambda$=0.5 more than doubles structural-break recall over Greedy (.13 vs. .05), while MMR$(0.7)$ leads on realized drawdown ($-1.22\%$ vs. $-0.87\%$), an observation we interpret as complementary regimes within the same Pareto frontier.
- **Solver computational cost** (App. I): SA \$0 / 45ms; QAOA-Rig 1 task = \$0.75 (Ankaa-3) or \$0.51 (Cepheus-1) at 500 shots.

The remaining narrative claims of this paper – that the formulation is solver-agnostic, the diversity gain is not an artifact of small $n$ or specific $K$, the policy generalizes across regimes, the formulation transfers across domains, and the gain over Greedy is not a tautological self-comparison – each rest on one of these appendix sections.

## 5. Related Work

**Risk in multi-agent systems.** Herding and crowding as systemic risk drivers are well established in behavioral economics (Banerjee, 1992; Khandani & Lo, 2011; Brown et al., 2022; Kang et al., 2021). The density-risk relationship has been documented empirically in financial markets (Brown et al., 2022; Kang et al., 2021), theoretically through herding models (Banerjee, 1992), and via fire-sale contagion mechanisms (Cont & Schaanning, 2017). We extend this line of work from prediction to decision.

**QUBO and combinatorial optimization.** QUBO provides a unified framework for NP-hard combinatorial problems (Glover et al., 2022; Lucas, 2014), with the underlying Ising formulation known to be NP-hard (Barahona, 1982). Applications include feature selection (Mücke et al., 2023), portfolio optimization (Mugel et al., 2022; Morapakula et al., 2025), and community detection (Negre et al., 2020), with recent work on QUBO encoding strategies (De Santis et al., 2026). Portfolio QUBO work optimizes returns, whereas we optimize monitoring diversity under a fragility-based risk model, a complementary decision problem that extends beyond finance to general MAS.

**Quantum optimization.** Quantum annealing (Kadowaki & Nishimori, 1998) and QAOA (Farhi et al., 2014; Zhou et al., 2020) provide QUBO-native solvers in the NISQ era (Preskill, 2018). Recent benchmarks show D-Wave hybrid solvers competitive with classical methods on dense instances (Kim et al., 2025), and quantum computing for finance is an active application area (Herman et al., 2023; Abbas et al., 2024). We contribute a concrete MAS application validated on real quantum hardware (Rigetti Ankaa-3 and Cepheus-1-108Q), with classical baselines for direct comparison and multi-seed hardware error bars.

## 6. Limitations and Future Work

**Problem scale.** Our controlled experiments use $n$=15 to enable exact verification, with scaling studies up to $n$=30 (SA) and $n$=20 (Rigetti QPU). The practical relevance of QUBO-based decision making and quantum backends emerges at larger scales ($n$>50), where exact enumeration is infeasible and classical heuristics may degrade. Quantum and hybrid approaches may provide advantages in such regimes.

**QAOA depth.** We use $p$=1 (single QAOA layer), which limits expressivity. Higher $p$ improves solution quality in theory (Zhou et al., 2020) but increases circuit depth and

noise sensitivity on current hardware.

**Static candidate pool.** The current formulation selects from a fixed pool. In practice, MAS monitoring requires rolling temporal windows in which the candidate pool and similarity structure evolve. Extending QUBO to warm-started sequential optimization is an open direction.

**Cross-domain validation depth.** Our deepest empirical analysis (lambda/K sweeps, scaling, learned policy, QPU runs) is on financial markets, where labeled crash horizons are available. The cross-domain results in Sec. 4 on inD traffic and SMAC v2 MARL are smaller-scale demonstrations of formulation transfer, not full-protocol replications; in particular, MARL is evaluated on 70 episodes. Operational-grade validation in those domains, with domain-specific labels, larger episode pools, and rolling candidate updates, is left to future work.

## 7. Conclusion

We have shown that fragility-aware monitoring in multi-agent systems can be formulated as a QUBO problem, providing a principled, solver-agnostic framework for discrete risk allocation. The formulation jointly optimizes risk coverage and diversity in a single quadratic objective without continuous relaxation. We validated it with four solvers, including multi-seed QAOA on Rigetti superconducting hardware (Ankaa-3 single-seed and Cepheus-1-108Q multi-seed; $9.83 total), all producing competitive solutions. The $\lambda$-parameterized risk-diversity tradeoff enables operators to balance monitoring breadth against risk concentration explicitly. These results establish QUBO as a natural computational abstraction for this class of MAS decision problems, with a clear path to quantum acceleration as hardware scales.

## Impact Statement

This work proposes a computational framework for risk monitoring in multi-agent systems. The primary intended application is systemic risk surveillance in financial markets, where improved diversity in monitoring coverage could help detect a broader range of failure modes. We do not foresee direct negative societal consequences from this methodology. The quantum computing component is validated at small scale and positioned as a forward-looking capability, not a recommendation for immediate deployment. All experiments use publicly available market data and cloud-accessible quantum hardware, ensuring reproducibility.

## Reproducibility

Code, data pipelines, and experiment scripts are in the supplementary material; classical solvers (Greedy, SA, Exact) need only `dwave-ocean-sdk` on any CPU, QAOA simulation needs `amazon-braket-sdk` (free), and the QPU runs need an AWS Braket account. Total QPU cost: $9.83 (18 tasks); recommended device for re-running is Rigetti Cepheus-1-108Q (Ankaa-3 was retired during this work).

## References


Abbas, A., Ambainis, A., Augustino, B., et al. Challenges and opportunities in quantum optimization. *Nature Reviews Physics*, 6:718–735, 2024. doi: 10.1038/s42254-024-00770-9.

Banerjee, A. V. A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817, 1992. doi: 10.2307/2118364.

Barahona, F. On the computational complexity of Ising spin glass models. *Journal of Physics A: Mathematical and General*, 15(10):3241–3253, 1982. doi: 10.1088/0305-4470/15/10/028.

Brown, G. W., Howard, P., and Lundblad, C. T. Crowded trades and tail risk. *The Review of Financial Studies*, 35 (7):3231–3271, 2022. doi: 10.1093/rfs/hhab107.

Chandola, V., Banerjee, A., and Kumar, V. Anomaly detection: A survey. *ACM Computing Surveys*, 41(3):1–58, 2009. doi: 10.1145/1541880.1541882.

Cont, R. and Schaanning, E. Fire sales, indirect contagion and systemic stress testing. *Norges Bank Working Paper*, (2017/2), 2017. doi: 10.2139/ssrn.2541114.

Cont, R. and Wagalath, L. Fire sales forensics: Measuring endogenous risk. *Mathematical Finance*, 26(4):835–866, 2016. doi: 10.1111/mafi.12071.

D-Wave Quantum Inc. D-Wave Ocean SDK. `https://docs.dwavequantum.com`, 2024. Version 9.x. Accessed March 2026.

De Santis, D., Tirone, S., Marmi, S., and Giovannetti, V. Optimized QUBO formulation methods for quantum computing. *Quantum Science and Technology*, 2026. doi: 10.1088/2058-9565/ae3b70.

Farhi, E., Goldstone, J., and Gutmann, S. A quantum approximate optimization algorithm. *arXiv preprint arXiv:1411.4028*, 2014.

Glover, F., Kochenberger, G., and Du, Y. Quantum bridge analytics I: A tutorial on formulating and using QUBO models. *Annals of Operations Research*, 314:141–183, 2022. doi: 10.1007/s10479-022-04634-2.

Greenwood, R., Landier, A., and Thesmar, D. Vulnerable banks. *Journal of Financial Economics*, 115(3):471–485, 2015. doi: 10.1016/j.jfineco.2014.11.006.

Herman, D., Googin, C., Liu, X., Sun, Y., Galda, A., et al. Quantum computing for finance. *Nature Reviews Physics*, 5:450–465, 2023. doi: 10.1038/s42254-023-00603-1.

Kadowaki, T. and Nishimori, H. Quantum annealing in the transverse Ising model. *Physical Review E*, 58(5): 5355–5363, 1998. doi: 10.1103/PhysRevE.58.5355.

Kang, W., Rouwenhorst, K. G., and Tang, K. Crowding and factor returns. *SSRN Electronic Journal*, 2021. doi: 10.2139/ssrn.3803954.

Khandani, A. E. and Lo, A. W. What happened to the quants in August 2007? Evidence from factors and transactions data. *Journal of Financial Markets*, 14(1):1–46, 2011. doi: 10.1016/j.finmar.2010.07.005.

Kim, S., Ahn, S. W., Suh, I. S., Dowling, A. W., Lee, E., and Luo, T. Quantum annealing for combinatorial optimization: A benchmarking study. *npj Quantum Information*, 11(1):77, 2025. doi: 10.1038/s41534-025-01020-1.

Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. Optimization by simulated annealing. *Science*, 220(4598): 671–680, 1983. doi: 10.1126/science.220.4598.671.

Lucas, A. Ising formulations of many NP problems. *Frontiers in Physics*, 2:5, 2014. doi: 10.3389/fphy.2014.00005.

Morapakula, S. N. V. et al. End-to-end portfolio optimization with hybrid quantum annealing. *Advanced Quantum Technologies*, 2025. doi: 10.1002/qute.202500753.

Mücke, S., Heese, R., Müller, S., Wolter, M., and Piatkowski, N. Feature selection on quantum computers. *Quantum Machine Intelligence*, 5(1):11, 2023. doi: 10.1007/s42484-023-00099-z.

Mugel, S., Kuchkovsky, C., Sánchez, E., Fernández-Lorenzo, S., Luis-Hita, J., Lizaso, E., and Orús, R. Dynamic portfolio optimization with real datasets using quantum processors and quantum-inspired tensor networks. *Physical Review Research*, 4(1):013006, 2022. doi: 10.1103/PhysRevResearch.4.013006.

Negre, C. F., Ushijima-Mwesigwa, H., and Mniszewski, S. M. Detecting multiple communities using quantum annealing on the D-Wave system. *PLoS ONE*, 15(2): e0227538, 2020. doi: 10.1371/journal.pone.0227538.

Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum*, 2:79, 2018. doi: 10.22331/q-2018-08-06-79.

Zhou, L., Wang, S.-T., Choi, S., Pichler, H., and Lukin, M. D. Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices. *Physical Review X*, 10:021067, 2020. doi: 10.1103/PhysRevX.10.021067.

## A. Density vs. Anomaly Baselines for Crash Prediction

This appendix supports the Crowding Paradox claim in Section 1: that systemic-risk-relevant states are *high-density* states, not low-density anomalies. On the same 2024-01-02 to 2026-02-02 financial test set used throughout the paper (523 trading days; train: 2017-2021), we compare GMM density and the derived Fragility Score against five standard anomaly-detection baselines. Crash labels are forward-looking $H$-day max drawdown beyond a $-3\sigma$ threshold; positive counts are 4 ($H$=5), 9 ($H$=10), and 24 ($H$=20).

*Table 4.* Crash-prediction AUROC and AUPRC for density-based scoring vs. anomaly-detection baselines on the financial test set. Density-based methods (top two rows) treat *high* probability as risk-relevant; anomaly methods (lower five) treat *low* probability as risk-relevant. The reversal in score direction is the Crowding Paradox.

| Method | $H$=5 d | | $H$=10 d | | $H$=20 d | |
|---|---|---|---|---|---|---|
| | AUROC | AUPRC | AUROC | AUPRC | AUROC | AUPRC |
| GMM density | **.941** | **.094** | **.943** | **.176** | **.961** | **.488** |
| Fragility Score | .941 | .096 | .942 | .167 | .960 | .460 |
| HMM ($k$=3) | .920 | .061 | .886 | .078 | .770 | .119 |
| Isolation Forest | .390 | .010 | .306 | .013 | .412 | .039 |
| OCSVM | .115 | .005 | .294 | .013 | .443 | .041 |
| LOF | .532 | .010 | .466 | .017 | .527 | .046 |
| PCA reconstruction | .802 | .023 | .715 | .033 | .606 | .057 |
| AE reconstruction | .431 | .009 | .364 | .014 | .535 | .061 |

GMM density and Fragility Score consistently outperform every anomaly baseline by a wide margin (AUROC $\geq$ .94 at all horizons; best anomaly baseline reaches .80 at $H$=5 and degrades thereafter). The dominance is even sharper on AUPRC, which is more informative for the rare-event regime. This empirically validates the central premise of the paper: density, not anomaly, is the operative risk signal for crowded multi-agent systems, and motivates using the density-derived Fragility Score as the QUBO reward.

## B. Risk-Diversity Pareto Front

*Table 5.* Full solver Pareto results at $n$=15, $K$=8, horizon=5d (extends Table 1 with all QAOA variants). SA/Exact: 5 data-resampling seeds. QAOA-sim: 5 classical-optimizer seeds. QAOA-Rig (Ankaa-3): single submission per $\lambda$ ($0.75/task; device retired during this work). QAOA-Rig (Cepheus-1-108Q): 3 independent QPU submissions per $\lambda$ ($0.51/task).

| $\lambda$ | Method | Avg FS | Diversity | $\lvert\mathcal{S}\rvert$ |
|---|---|---|---|---|
| 0.0 | Greedy | .367$\pm$.002 | .476$\pm$.035 | 8.0 |
| | SA | .366$\pm$.002 | .487$\pm$.029 | 8.0 |
| | Exact | .367$\pm$.002 | .476$\pm$.035 | 8.0 |
| | QAOA-sim | .367$\pm$.000 | .506$\pm$.057 | 8.0 |
| | QAOA-Rig (Ankaa-3) | .368 | .481 | 8 |
| 0.5 | Greedy | .367$\pm$.002 | .476$\pm$.035 | 8.0 |
| | SA | .348$\pm$.025 | .591$\pm$.070 | 8.0 |
| | Exact | .348$\pm$.025 | .591$\pm$.070 | 8.0 |
| | QAOA-sim | .366$\pm$.001 | .586$\pm$.027 | 8.0 |
| | QAOA-Rig (Ankaa-3) | .364 | .611 | 8 |
| | QAOA-Rig (Cepheus-1) | .366$\pm$.001 | .592$\pm$.020 | 8.0 |
| 1.0 | Greedy | .367$\pm$.002 | .476$\pm$.035 | 8.0 |
| | SA | .337$\pm$.032 | .639$\pm$.062 | 7.2 |
| | Exact | .337$\pm$.032 | .639$\pm$.062 | 7.2 |
| | QAOA-sim | .365$\pm$.002 | .628$\pm$.016 | 7.0 |
| | QAOA-Rig (Ankaa-3) | .362 | .640 | 7 |
| | QAOA-Rig (Cepheus-1) | .367$\pm$.001 | .616$\pm$.007 | 7.0 |

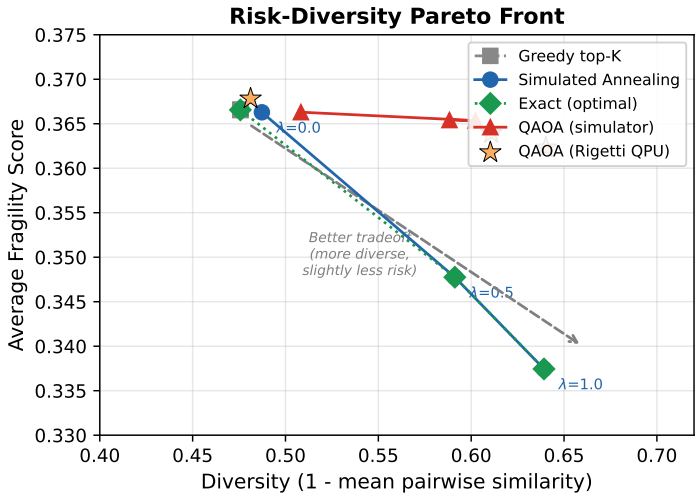


*Figure 2.* Risk-diversity Pareto front across $\lambda \in \{0.0, 0.5, 1.0\}$ at $n$=15, $K$=8, horizon=5d. Greedy (gray) is fixed regardless of $\lambda$. SA and Exact trace the optimal Pareto curve; QAOA on Rigetti (Ankaa-3 single-seed and Cepheus-1-108Q multi-seed) tracks the frontier within hardware noise.

Greedy selection is Pareto-dominated: QUBO-aware methods achieve higher diversity with competitive Avg FS. The $\lambda$ parameter provides smooth, interpretable control along the frontier. QAOA solutions, both simulated and hardware, lie close to the SA/Exact frontier; the multi-seed Cepheus-1-108Q runs (FS $.366 \pm .001$ / Div $.592 \pm .020$ at $\lambda$=0.5, FS $.367 \pm .001$ / Div $.616 \pm .007$ at $\lambda$=1.0) sit slightly inside the Exact frontier as expected from the QAOA $p$=1 approximation plus 500-shot measurement noise.

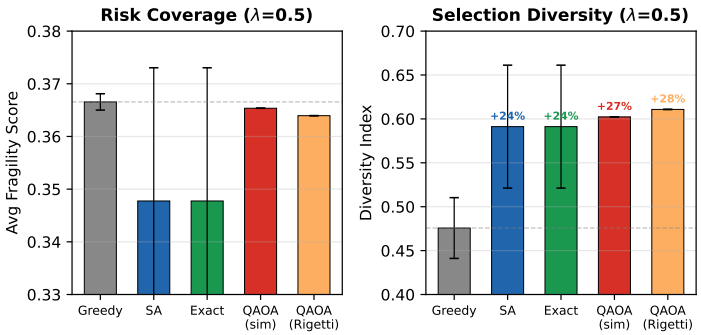


*Figure 3.* Solver comparison at $\lambda$=0.5. Left: Avg Fragility Score. Right: Diversity. Percentages denote improvement over the Greedy baseline. Error bars: $\pm 1$ std across 5 seeds.

**Solver comparison at $\lambda$=0.5.** SA and Exact achieve identical solutions ($+24\%$ diversity over Greedy). QAOA-simulator and QAOA-Rig at $\lambda$=0.5 achieve $+24\%$ to $+28\%$ diversity respectively, while maintaining FS within $1\%$ of the Greedy baseline.

## C. Cross-Horizon Consistency

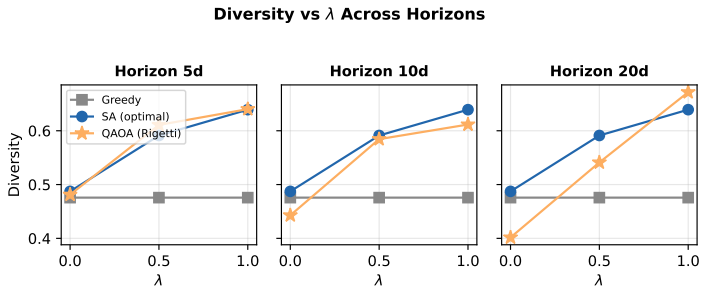


*Figure 4.* Diversity as a function of $\lambda$ across crash horizons (5d, 10d, 20d). SA consistently increases diversity with $\lambda$; Rigetti results follow the same trend with higher variance.

The QUBO objective depends on FS and similarity rather than on horizon-specific labels, so the optimization is horizon-agnostic by construction. The variation in QAOA-Rig results across horizons reflects finite sampling (500 shots) and hardware noise rather than systematic horizon dependence.

*Table 6.* Cross-horizon results at $\lambda$=0.5 for Greedy, SA, and QAOA-Rig (Ankaa-3 single-seed).

| Horizon | Method | Avg FS | Diversity | $\lvert\mathcal{S}\rvert$ |
|---|---|---|---|---|
| 5d | Greedy | .367 | .476 | 8 |
| | SA/Exact | .348 | .591 | 8 |
| | QAOA-Rig | .364 | .611 | 8 |
| 10d | Greedy | .367 | .476 | 8 |
| | SA/Exact | .348 | .591 | 8 |
| | QAOA-Rig | .366 | .584 | 8 |
| 20d | Greedy | .367 | .476 | 8 |
| | SA/Exact | .348 | .591 | 8 |
| | QAOA-Rig | .367 | .541 | 8 |

## D. Scaling Behavior ($n$=15 to 30)

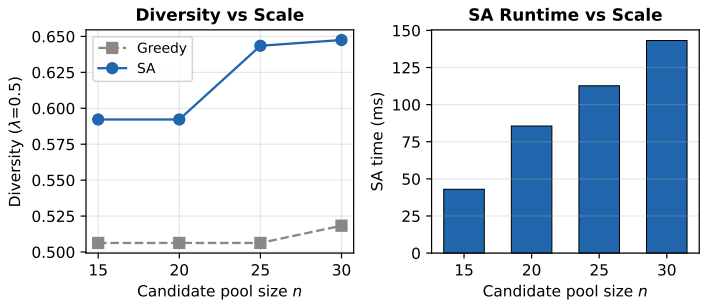


*Figure 5.* Left: SA diversity gain over Greedy remains stable as the candidate pool grows from $n$=15 to 30 ($\lambda$=0.5). Right: SA runtime scales from 43ms to 143ms.

The diversity advantage of QUBO-based optimization is not an artifact of small $n$. As the candidate pool grows from 15 to 30, SA consistently achieves diversity ∼0.59–0.65 at $\lambda$=0.5, compared with Greedy's ∼0.51–0.52. SA runtime remains practical, growing from 43ms ($n$=15) to 143ms ($n$=30). We also executed QAOA on Rigetti Ankaa-3 at $n$=20 (630-gate circuit, \$0.75/task), obtaining diversity 0.577 at $\lambda$=0.5, confirming that the formulation remains executable on real quantum hardware beyond the smallest verifiable scale.

**Extended scaling to $n$=200.** To verify that the QUBO advantage is not an artifact of small problem size, we scale the candidate pool to $n \in \{50, 100, 200\}$ on the full financial test set. At $n$=200 with $\lambda$=0.5, SA achieves diversity 0.851 compared with Greedy's 0.514, a +66% improvement. The diversity gain grows monotonically with $n$: +24% at $n$=15, +31% at $n$=50, +46% at $n$=100, and +66% at $n$=200. SA runtime grows from 44ms ($n$=15) to 56 seconds ($n$=200), remaining practical for daily monitoring.

## E. Sensitivity to $K$ and $\lambda$

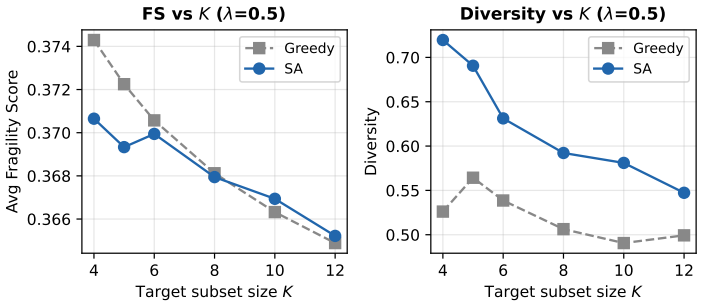


*Figure 6.* Diversity and FS as a function of target subset size $K$ ($n$=15, $\lambda$=0.5). SA outperforms Greedy in diversity across all $K$ values.

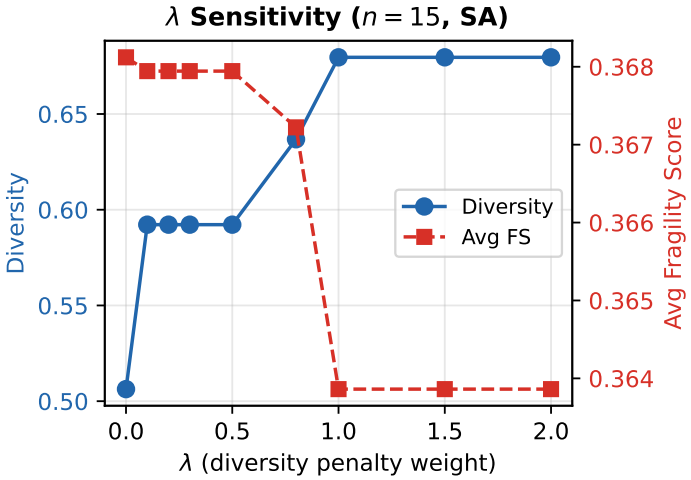


*Figure 7.* Fine-grained $\lambda$ sweep ($n$=15, SA). Diversity (blue) exhibits a sharp onset near $\lambda \approx 0.1$ and saturates beyond $\lambda \approx 1.0$. FS (red) decreases monotonically but remains within 1% of baseline for $\lambda \leq 0.5$.

The SA-over-Greedy diversity advantage holds across target subset sizes $K \in \{4, 5, 6, 8, 10, 12\}$, confirming that the result is not specific to $K$=8. The gain is most pronounced at smaller $K$, where selecting few diverse states is combinatorially hardest. The $\lambda$ sweep shows a sharp diversity onset near $\lambda \approx 0.1$ – where the similarity penalty first becomes strong enough to alter the selection – followed by saturation beyond $\lambda \approx 1.0$. FS decreases monotonically with $\lambda$ but remains within 1% of the Greedy baseline for $\lambda \leq 0.5$, indicating a favorable operating region. Pilot experiments with cosine similarity yield qualitatively similar behavior, suggesting that the effect is not kernel-specific.

## F. Learned $\lambda$ Policy

A fixed $\lambda$ ignores the structure of each candidate pool. We train a lightweight MLP policy (10-dim pool statistics $\rightarrow \lambda$) via REINFORCE, optimizing a composite reward of FS, diversity, and cardinality satisfaction.

*Table 7.* Learned $\lambda$ on held-out test episodes. Objective is the composite reward used during training.

| Method | Objective | FS | Diversity | $\hat{\lambda}$ |
|---|---|---|---|---|
| Greedy | 0.908 | .406 | .502 | 0 |
| Fixed $\lambda$=0.5 | 1.055 | .396 | .660 | 0.50 |
| Learned $\lambda$ | **1.071** | .394 | .677 | 0.64 |
| Grid (oracle) | 1.076 | .393 | .684 | 0.78 |

The learned policy achieves $99.5\%$ of the oracle grid-search objective on held-out test data, generalizing from the training period (pre-2022) to the test period (post-2024) without retraining.

## G. Cross-Domain Generalization

*Table 8.* Cross-domain QUBO results at $\lambda$=0.5, $K$=8, $n$=50, 3 seeds. Diversity is $1 - \bar{S}$ on the domain-specific RBF kernel; absolute values are not comparable across domains because the kernel bandwidth differs.

| Domain | Greedy Div | SA Div | Relative gain |
|---|---|---|---|
| Financial | .518 | .657$\pm$.028 | +27% |
| Traffic (inD)$^{\dagger}$ | .000 | .421$\pm$.012 | saturated |
| MARL (SMAC v2) | .763 | .816$\pm$.009 | +7% |

$^{\dagger}$Traffic Greedy diversity rounds to .000 (.0001 before rounding) because the top-FS candidate pool consists of consecutive frames at the same crowded intersection, which RBF-saturate to similarity $\approx 1$ at the chosen bandwidth; SA spreads selections across distinct configurations and recovers measurable diversity.

The QUBO formulation provides positive diversity gains over Greedy in all three domains, but the magnitude varies: large in finance and traffic, modest in MARL where Greedy already selects relatively diverse episodes (.763). The Traffic Greedy diversity rounding to zero is a genuine RBF saturation effect, not a measurement bug, and is itself the strongest case for moving from score-and-threshold to QUBO subset selection: Greedy returns a degenerate redundant subset, while SA returns a useful one.

## H. Downstream Stress Capture

To break the tautological coupling between the QUBO objective (similarity penalty) and the in-paper diversity metric (1 minus mean similarity), we evaluate each method on *forward-looking* market metrics that are independent of the optimization signal: realized 5-day forward drawdown of SPY at the selected dates, realized 5-day forward volatility, temporal coverage (number of selections that are at least 7 trading days apart), and recall on the labeled `structural_break` events at a 20-day window.

*Table 9.* Downstream stress-capture metrics on the financial test set ($n$=15, $K$=8, 5 seeds; mean$\pm$std). Forward drawdown is the realized 5-day max drawdown of SPY at the selected dates (more negative = stress days successfully captured); recall is computed against the 73 labeled structural-break events.

| Method | Fwd DD-5d (%) | Fwd vol-5d (%) | Cov. 7d | Recall (SB,20d) |
|---|---|---|---|---|
| Greedy | $-0.90\pm.13$ | $0.50\pm.06$ | 5.4 | .05 |
| SA, $\lambda$=0.5 | $-0.87\pm.06$ | $0.56\pm.04$ | 5.6 | **.13** |
| SA, $\lambda$=1.0 | $-0.70\pm.19$ | $0.54\pm.06$ | 5.4 | .10 |
| MMR(0.5) | $-1.17\pm.17$ | $0.65\pm.05$ | 4.8 | .10 |
| MMR(0.7) | $\mathbf{-1.22}\pm.22$ | $0.64\pm.06$ | 5.2 | **.15** |
| $k$-DPP | $-1.01\pm.27$ | $0.61\pm.05$ | 5.0 | .08 |

The downstream metrics support the framing of QUBO as a controllable frontier rather than as a single operating point. At $\lambda$=0.5, QUBO-SA more than doubles the structural-break recall over Greedy (.13 versus .05) while preserving comparable

temporal coverage, and at higher $\lambda$ the same formulation reaches a realized forward drawdown of $-0.70\%$. MMR at diversity weight $0.7$ achieves $-1.22\%$ realized drawdown, which is a strong outcome but is obtained at a single fixed operating point, because MMR exposes only its diversity weight $\mu$ and reaching a different stress regime would require switching to another method. The QUBO formulation, in contrast, dials the same property through $\lambda$ along a continuous Pareto curve that contains MMR and $k$-DPP as fixed points; once a deployment specifies which downstream metric matters (recall, drawdown, coverage, or a weighted combination), the appropriate $\lambda$ is read off the frontier rather than chosen between methods. The claim our experiments therefore support is one of generality across operating points, rather than of point-by-point dominance over any individual baseline.

## I. Solver Computational Cost

*Table 10.* Solver cost comparison per QUBO instance ($n$=15).

| Solver | Time | Hardware | USD Cost |
|---|---|---|---|
| Greedy | <1ms | CPU | \$0 |
| SA (200 reads) | 45ms | CPU | \$0 |
| Exact | 26ms | CPU | \$0 |
| QAOA-sim | ∼7s | CPU | \$0 |
| QAOA-Rig (Ankaa-3) | ∼2min* | QPU | \$0.75 |
| QAOA-Rig (Cepheus-1) | ∼40s* | QPU | \$0.51 |

*Includes local param. optimization + QPU queue + 500 shots.

## J. Why QUBO over Gradient Methods, and QAOA as Forward-Looking Backend

**Why QUBO over gradient methods?** Subset selection is a discrete problem: the decision variables $z_i \in \{0, 1\}$ do not admit continuous relaxation without introducing approximation. Continuous relaxation methods such as Gumbel-Softmax or straight-through estimators struggle to enforce exact cardinality and pairwise diversity constraints simultaneously, and typically require heuristic rounding that breaks optimality guarantees. QUBO encodes the entire objective, including the cardinality constraint via the penalty term $(\sum_i z_i - K)^2$, as a single polynomial (Glover et al., 2022). The same instance can be dispatched to any compatible solver without reformulation, from classical SA to quantum annealers to gate-model QAOA.

**QAOA as a forward-looking backend.** The goal of this work is not to demonstrate quantum advantage at small scale but to verify that the same decision formulation executes on real quantum hardware without modification. At $n$=15, classical SA finds the exact optimum in 45ms on a single CPU core; quantum advantage is not expected at this scale. We verify *computational compatibility*: the QUBO formulation provides a natural interface to quantum hardware via the standard QAOA ansatz, requiring no problem-specific circuit design. The multi-seed Cepheus-1-108Q evaluation (Sec. 4) shows the per-seed std at $\lambda$=1.0 is .007, while the gap between the multi-seed mean (.616) and Exact (.639) is .023, suggesting the residual gap is dominated by the $p$=1 ansatz approximation rather than measurement noise; raising $p$ on simulator and on hardware (subject to coherence-time budgets) is the natural next experiment. We further extend QAOA execution to $n$=20 on Rigetti Ankaa-3 (630-gate circuit), confirming the formulation remains executable beyond the smallest verifiable scale. As quantum hardware scales, larger QUBO instances from denser MAS monitoring scenarios could benefit from quantum parallelism, while the problem formulation and evaluation framework remain unchanged.