Parlay Identification of Ising Couplings in Correlated Binary Event Markets

Raeez Lorgat

April 2025

Abstract

Prediction markets on correlated binary events produce joint distributions that independent marginal pricing cannot represent. We study an event-outcome Ising model in which each binary event is a ± 1 spin, the marginal $\mathbb{P}(\omega_i = +1) = p_i$ maps to an external field $h_i = \frac{1}{2}\operatorname{logit}(p_i)$, and the pairwise interaction term J_{ij} of the Boltzmann distribution encodes the excess joint probability of the two events beyond their marginals. A parlay contract on the pair (i, j) resolves at $(\omega_i, \omega_j) \in \{-1, +1\}^2$; we observe N_p such resolutions and estimate J_{ij} by pair-marginal maximum likelihood on the empirical joint-outcome counts $(N_{++}, N_{+-}, N_{-+}, N_{-})$. This is the two-site log-linear exponential family, and so situates in the classical auto-logistic framework of Besag [27] and the Boltzmann-machine formulation of Ackley, Hinton, and Sejnowski [28]; our contribution is an explicit admissible-parameter characterisation, Schur-complement efficient variance, non-asymptotic concentration, and embedding-bias expansion in the prediction-market setting. Four results follow. First, a *closed-form identifiability* statement: for $M = 2$ events, the triple (p_i, p_j, π_{ij}) of two marginals and the joint-outcome frequency determines (h_i, h_j, J_{ij}) uniquely, with an explicit admissible region (Theorem 3.2). Second, an *efficient MLE* for J_{ij} from N_p resolved parlay trades: the two-site Ising model is a regular three-parameter exponential family with sufficient statistics $(\omega_i, \omega_j, \omega_i\omega_j)$; we derive the 3×3 Fisher information matrix, the Schur-complement efficient variance, and prove consistency and asymptotic normality (Theorem 3.6). Third, a *non-asymptotic concentration bound*: $|\hat{J}_{ij} - J^*_{ij}| \le \epsilon$ with probability $1 - \delta$ for $N_p \ge C(\theta^*)\,\epsilon^{-2}\log(2/\delta)$, for ϵ in a local neighbourhood of θ^* and with $C(\theta^*)$ computed explicitly from the two-site moments (Theorem 4.1). Fourth, a *sharpened bias bound* for the two-site estimator applied inside an $M > 2$ network: under a Plefka / TAP expansion the leading embedding bias is $\Delta_{ij} = -\frac{1}{2}\sum_{k \ne i,j} J^*_{ik}J^*_{jk}(1 - m_k^{*2}) + O(J^3_{\max})$ (Theorem 5.1). A *robustness* result bounds the perturbation of inferred (h_i, J_{ij}) under an L^∞ perturbation of the quoted marginals (Theorem 6.1).

Novelty statement. The Ising model [6], the Fisher z-transform [5], the auto-logistic reparameterisation [27], and the Boltzmann-machine exponential-family formulation [28] are classical. Applications of Ising-type models to financial markets [8, 10, 12, 13, 14] use spins as *trader decisions* and interpret J_{ij} as market-participant imitation strength. The model here is different: spins are *event outcomes* and J_{ij} is identified from the empirical frequency of joint resolutions of parlay contracts. The ℓ_1-regularised Ising structure recovery of Ravikumar-Wainwright-Lafferty [18], the information-theoretic lower bound of Santhanam-Wainwright [29], the efficient structure-learning algorithm of Bresler [30], and the interaction-screening estimator of Vuffray et al. [19] address the large-M simultaneous-selection problem from samples of the full spin vector. Our contribution is complementary: pair-marginal samples rather than full-vector samples, unpenalised estimation of a single edge rather than sparsity-regularised recovery of the full adjacency, and fully parametric efficiency rather than sample-optimal structure recovery. We establish (1) an explicit admissible-parameter region for the two-site Ising exponential family, (2) the Schur-complement efficient Fisher bound on the J coordinate with an explicit correction in the off-diagonal regime $m_i m_j \neq 0$, (3) a Hoeffding-route non-asymptotic bound local to θ^*, and (4) a Plefka / TAP leading-order bias expansion [32, 33] for the two-site estimator applied to a marginal of an M-site network.

Prior art positioning. The closest work on what market prices do and do not identify probabilistically is Manski [31], which showed that under heterogeneous beliefs or risk aversion a prediction-market price is a possibly biased estimator of a specific functional of the cross-trader belief distribution rather than the objective probability. We adopt the standard identification $\mathbb{P}(\omega_i = +1) = p_i$ conditional on a risk-neutral representative trader at the scoring-rule equilibrium; Section 2 makes this explicit and Section 6 bounds the consequences of deviation. The closest work on parlay markets specifically is Rana et al. [3], which established that J_{ij} in the Ising parameterisation cannot be recovered from static marginal prices alone and argued that parlay trades are structurally required. The combinatorial prediction-market literature [1, 16, 17] prices joint-outcome contracts when the joint distribution is given; our results provide one route to estimating that distribution from parlay-outcome data under the pairwise-Ising parameterisation.

Contents

1 Introduction

1.1 The correlation problem for binary event markets

Consider two binary prediction markets. Market i trades contracts paying \$1 if event e_i ("Team A wins the championship") occurs, and zero otherwise. Market j trades contracts on e_j ("Team B retains the title"). Under standard market-scoring-rule pricing [2], the marginal prices $p_i, p_j \in (0, 1)$ are the markets' best estimates of $\mathbb{P}(\omega_i = +1)$ and $\mathbb{P}(\omega_j = +1)$ respectively.

If the events are independent, the joint probability is $p_i p_j$. Two events resolved by separate tournaments on different continents can be modelled this way.

Many events are not independent. If Team A and Team B compete in the same league, a title change for Team A shifts Team B's probabilities through scheduling, seeding, and competitive-balance effects. In that case the independence-implied joint probability $p_i p_j$ understates or overstates the true joint by an amount that depends on the pairwise correlation. Log-linear aggregation [2, 4], which assumes conditional independence of signals given a common fundamental, is the zero-correlation limit of any model with a nontrivial joint distribution over outcomes.

1.2 The Ising model for binary event markets

Several parameterisations of the joint distribution are in use: rolling sample correlations, factor models, copulas, Bayesian networks. The parameterisation studied here is the pairwise binary maximum-entropy model, equivalently the Ising model: M binary events $e_1, \dots, e_M$ with outcomes $\omega_i \in \{-1, +1\}$, joint distribution

$$\mathbb{P}(\omega_1, \dots, \omega_M) = \frac{1}{Z} \exp\left(\sum_i h_i \omega_i + \sum_{i<j} J_{ij} \omega_i \omega_j \right), \tag{1}$$

with h_i the external field on site i and J_{ij} the pairwise interaction. Three structural advantages motivate this choice:

1. The Ising density (1) is the pairwise maximum-entropy distribution on $\{-1, +1\}^M$ with specified marginals and pairwise correlations; among distributions matching those moments it assumes the least additional structure.
2. The parameters (h_i, J_{ij}) have direct interpretations as logit-marginals and as pairwise log-likelihood-ratio corrections for joint outcomes; they are communicable to traders and regulators.
3. The pair (h, J) determines a consistent joint distribution over all 2^M outcomes, enabling the pricing of parlays and other joint-outcome contracts.

These advantages come at explicit costs. The pairwise maximum-entropy form cannot represent higher-order dependencies such as three-event conjunctions with joint correlation and pairwise independence, asymmetric tail dependence, or non-linear dependence. Copulas, vine copulas, factor models, and directed graphical models are each more expressive on some axes and less on others. This paper derives identifiability and estimation results for the Ising model; the results state what one can learn from market data under the Ising parameterisation, not that the parameterisation is universally correct.

1.3 What this paper does

The paper focuses on a single well-defined question: given marginal prices p_i and N_p resolved parlay contracts on pair (i, j) reporting joint outcomes (ω_i, ω_j), what can be inferred about the Ising parameters (h_i, J_{ij}), and with what finite-sample guarantees?

(1) Identifiability from pair-level outcome frequencies (Section 2). For $M = 2$, the triple (p_i, p_j, π_{ij}) of two marginals and the empirical joint-outcome frequency $\pi_{ij} = N_{++}/N_p$ determines (h_i, h_j, J_{ij}) uniquely, with an explicit characterisation of the admissible region (Theorem 3.2). For $M > 2$, the same closed-form inversion applied to a pair inside a larger network has a computable embedding bias; the leading-order form is given in Proposition 3.1 and sharpened in Theorem 5.1.

(2) Efficient MLE (Section 3.3). The two-site Ising model is a regular three-parameter exponential family with sufficient statistics $(\omega_i, \omega_j, \omega_i\omega_j)$. We give the log-likelihood, derive the 3×3 Fisher information matrix, and use the Schur complement to express the efficient asymptotic variance of $\hat{J}_{ij}$ (Theorem 3.6); the Schur correction is a constant-factor improvement over the marginal bound $1/[N_p(1 - C_{ij}^{*2})]$ and is nonzero whenever the marginals are biased ($m_i m_j \neq 0$), which is the generic case.

(3) Non-asymptotic concentration bound (Section 4). For any target (ϵ, δ) with ϵ in a local neighbourhood of θ^*, we give a sample size N_p such that $|\hat{J}_{ij} - J_{ij}^*| \leq \epsilon$ with probability $\geq 1 - \delta$, with the constants computed from the two-site moments at the truth (Theorem 4.1). The rate $\epsilon^{-2} \log(1/\delta)$ matches the parametric MLE rate; the information-theoretic lower bounds of [29] apply to simultaneous structure recovery and are not directly comparable to the pair-level question studied here.

(4) Sharpened $M > 2$ embedding bias (Section 5). When the two-site closed-form estimator is applied to a pair (i, j) inside a network of $M > 2$ coupled events, the resulting estimator is biased. Theorem 5.1 gives the leading bias $\Delta_{ij} = -\frac{1}{2} \sum_{k \neq i,j} J_{ik}^* J_{jk}^* (1 - m_k^{*2}) + O(J_{\max}^3)$ via a Plefka / TAP expansion [32, 33], and states a sufficient condition (small $J_{\max}$, sparse network) under which $|\Delta_{ij}| \ll |J_{ij}^*|$.

(5) Robustness to quoted-marginal noise (Section 6). Market-scoring-rule prices are perturbed by inventory and liquidity effects away from the equilibrium that identifies them with the Ising marginals. Theorem 6.1 bounds the propagation of an L^∞ perturbation in (p_i, p_j, π_{ij}) to the inferred (h_i, J_{ij}); the perturbation is treated exogenously.

(6) Fortification proofs (Section 7). A self-contained proof of consistency via Wald's theorem and of the Schur-complement form of the efficient Cramér-Rao bound, with numerical illustration of the Schur correction in a representative off-diagonal regime.

Sections 2.3 and 2.4 record a Lipschitz bound on the mean-field iteration for the Ising density (1), which is all the identification theorems require; phase-transition structure is not load-bearing and is noted only to locate the regime where the two-site Fisher information is positive-definite.

1.4 What this paper does *not* do

To scope the claims precisely:

- No empirical calibration of $\hat{J}$ against live market data is presented; the concentration bound is a finite-sample statement, not a measurement.
- No claim that the Ising parameterisation is more faithful than copulas, factor models, or directed graphical models in any specific domain; the paper states what follows *under* the Ising parameterisation.
- No mechanism-design claim about implementing these estimators on a specific exchange architecture; the results are statements about estimators on observations, independent of the settlement layer.

- A leading-order bridge from marginal-market lead-lag statistics (Granger causality) to J_{ij} was explored in earlier drafts; it is relegated to a future-work remark (Section 8.3) because the bridge carries second-order corrections whose sign and magnitude are sensitive to microstructure, making it less suitable for theorem-weight claims than the parlay-based MLE.

2 The Ising Model for Binary Event Markets

2.1 Boltzmann distribution, fields, and couplings

Consider M binary events $e_1, \dots, e_M$ with outcomes $\omega_i \in \{-1, +1\}$; $\omega_i = +1$ denotes e_i occurring.

Definition 2.1 (Ising Hamiltonian)**.** The Ising Hamiltonian for M events is

$$H(\omega) \;=\; -\sum_{i=1}^{M} h_i \omega_i \;-\; \sum_{1 \le i < j \le M} J_{ij}\, \omega_i \omega_j, \tag{2}$$

where $h_i \in \mathbb{R}$ is the external field on site i and $J_{ij} \in \mathbb{R}$ is the pairwise interaction; $J_{ij} > 0$ means the events e_i, e_j tend to co-occur, $J_{ij} < 0$ means they tend to anti-correlate.

The joint distribution is the Boltzmann density

$$\mathbb{P}(\omega_1, \dots, \omega_M) \;=\; \frac{1}{Z} \exp(-H(\omega)) \;=\; \frac{1}{Z} \exp\left(\sum_i h_i \omega_i + \sum_{i<j} J_{ij} \omega_i \omega_j \right), \tag{3}$$

with partition function $Z = \sum_{\omega \in \{-1,+1\}^M} \exp(-H(\omega))$.

Remark 2.2 (Mapping to marginal prices)**.** If the market on event e_i quotes marginal price p_i and that price identifies the Ising marginal $\mathbb{P}(\omega_i = +1) = p_i$, then the field takes the half-logit form

$$h_i \;=\; \tfrac{1}{2} \operatorname{logit}(p_i) \;=\; \tfrac{1}{2} \ln \frac{p_i}{1 - p_i}. \tag{4}$$

The factor of $1/2$ arises because the Ising convention uses $\omega_i \in \{-1, +1\}$ rather than $\{0, 1\}$. Under the $\{0, 1\}$ convention with $\sigma_i = (\omega_i + 1)/2$ the field becomes $\tilde{h}_i = \operatorname{logit}(p_i)$ and $\tilde{J}_{ij} = 4J_{ij}$, recovering the logistic parameterisation. The identification $\mathbb{P}(\omega_i = +1) = p_i$ requires that the market be at the scoring-rule equilibrium for its liquidity parameter; inventory and liquidity perturbations from that equilibrium propagate an error into h_i of magnitude $O(1/b)$ for liquidity parameter b (addressed quantitatively in Section 6).

2.2 Marginals and exact computation for small M

The marginal probability of event e_k occurring is

$$\mathbb{P}(\omega_k = +1) \;=\; \frac{\sum_{\omega : \omega_k = +1} \exp(-H(\omega))}{Z} \;=\; \frac{Z_k^+}{Z}, \tag{5}$$

where Z_k^+ sums over configurations with $\omega_k = +1$.

Proposition 2.3 (Exact marginals for small M)**.** *For $M \le 5$ events with arbitrary interaction matrix J, the marginal probability $\mathbb{P}(\omega_k = +1)$ and the partition function Z are computable by direct enumeration of $2^M \le 32$ configurations at cost $O(M^2 \cdot 2^M)$. For $M = 5$: 32 configurations, each requiring at most 15 multiplications, total ≤ 480 operations.*

Proof. By enumeration. Each configuration $\omega \in \{-1, +1\}^M$ contributes $\exp(-H(\omega))$ to Z and to either Z_k^+ or Z_k^-. Evaluating $H(\omega)$ requires M field terms and $\binom{M}{2}$ interaction terms, i.e. $O(M^2)$ operations per configuration. □

2.3 Mean-field approximation for larger M

For $M > 5$ the direct sum becomes expensive ($2^{20} \approx 10^6$, $2^{50} \approx 10^{15}$). The mean-field approximation replaces the joint Boltzmann distribution with a product of independent distributions minimising the KL divergence to the true distribution.

Definition 2.4 (Mean-field equations)**.** The mean-field magnetisations $\hat{m}_i = \mathbb{E}[\omega_i]$ satisfy

$$\hat{m}_i \;=\; \tanh\left(h_i + \sum_{j\neq i} J_{ij}\hat{m}_j\right), \qquad i = 1, \ldots, M. \tag{6}$$

The mean-field marginal is $\hat{p}_i = (1 + \hat{m}_i)/2$.

Theorem 2.5 (Mean-field convergence)**.** *Let $J_{\max} = \max_{i,j} |J_{ij}|$ and $d_{\max} = \max_i |\{j : J_{ij} \neq 0\}|$ be the maximum degree of the interaction graph.*

(i) If $d_{\max} \cdot J_{\max} < 1$, the mean-field iteration (6) converges to a unique fixed point from any initial condition, at rate $r_{\mathrm{MF}} = d_{\max} J_{\max}$.

(ii) The approximation error for marginal expectations satisfies $|\hat{m}_i - m_i| \leq d_{\max} J_{\max}^2/(1 - d_{\max} J_{\max})$, where $m_i = \mathbb{E}_{\mathbb{P}}[\omega_i]$ under the true Boltzmann distribution.

(iii) The mean-field free energy satisfies $F_{\mathrm{MF}} \geq -\ln Z$.

(iv) The iteration converges to additive error ε in $O(M d_{\max} \lceil \log(1/\varepsilon)/\log(1/r_{\mathrm{MF}}) \rceil)$ operations.

Proof. Part (i): Define $\Psi : \mathbb{R}^M \to \mathbb{R}^M$ by $\Psi_i(m) = \tanh(h_i + \sum_j J_{ij} m_j)$. The Jacobian entries are $\partial \Psi_i/\partial m_j = J_{ij}(1 - \Psi_i^2)$. Since $|1 - \tanh^2(x)| \leq 1$, the row-sum bound gives $\|\nabla \Psi\|_\infty \leq d_{\max} J_{\max}$. For $d_{\max} J_{\max} < 1$, Ψ is a contraction on $[-1, 1]^M$ with Lipschitz constant r_{MF}; Banach's theorem gives the unique fixed point.

Part (ii): The Gibbs variational principle states that the mean-field product distribution minimises $\mathrm{KL}(q\|p)$. A perturbation expansion gives a first-order correction that vanishes (the mean-field equations are stationarity conditions) and a second-order correction $O(J_{\max}^2)$ per interaction; summing over the $d_{\max}$ interactions at each site and applying the contraction factor gives the stated bound.

Part (iii): Bogoliubov's inequality (Jensen applied to $-\ln$) gives $-\ln Z \leq \mathbb{E}_q[-H] + \sum_i \mathbb{E}_q[\ln q_i] = F_{\mathrm{MF}}$.

Part (iv): Each iteration costs $M d_{\max}$ multiplications; convergence to ε requires $\lceil \log(1/\varepsilon)/\log(1/r_{\mathrm{MF}}) \rceil$ iterations. □

2.4 Instability and linear susceptibility

The Ising pricing iteration admits a canonical instability at $\|J\|_{\mathrm{op}} = 1$: beyond this threshold the fixed point $\hat{m} = 0$ ceases to be linearly stable. This is stated below in the specific form needed by the estimators of Sections 3-6, which require that the pair (i, j) operate in a regime where the two-site distribution is well-defined and the MLE Fisher information is nonsingular.

Definition 2.6 (Gershgorin coupling bound)**.** For the interaction matrix J, define

$$\beta_{\mathrm{eff}}(J) \;:=\; d_{\max} J_{\max}. \tag{7}$$

This is the Gershgorin / maximum-absolute-row-sum bound on $\|J\|_{\mathrm{op}}$; it is an upper bound on the spectral radius, not the spectral radius itself. In general $\rho(J) \leq \|J\|_{\mathrm{op}} \leq \beta_{\mathrm{eff}}(J)$ with strict inequality.

Theorem 2.7 (Mean-field instability threshold)**.** *Let J satisfy $\beta_{\text{eff}}(J) < 1$ and consider the zero-field limit $h = 0$.*

(i) The zero-magnetisation fixed point $\hat{m} = 0$ is linearly stable iff $\|J\|_{\text{op}} < 1$. The sharp instability threshold is $\|J\|_{\text{op}} = 1$.

(ii) If in addition J is irreducible with non-negative entries, the bifurcation at $\|J\|_{\text{op}} = 1$ is along the Perron eigenvector, to two symmetric fixed points $\pm\hat{m}^* \neq 0$.

(iii) Expanding the fixed-point equation near the transition gives $|\hat{m}^*| \sim (\|J\|_{\text{op}} - 1)^{1/2}$ within the mean-field approximation (i.e. the mean-field critical exponent $\beta_{\text{crit}} = 1/2$).

Proof. Part (i): At $h = 0$, $\hat{m} = 0$ solves (6) since $\tanh(0) = 0$. The Jacobian of Ψ at this fixed point is J (using $\tanh'(0) = 1$). Linear stability of a fixed point of an iterated map is equivalent to $\rho(J) < 1$, which for symmetric J equals $\|J\|_{\text{op}}$.

Part (ii): Under irreducibility and non-negativity, Perron-Frobenius gives a unique largest eigenvalue $\mu_1 = \|J\|_{\text{op}}$ with strictly positive eigenvector v_1. For $\mu_1 > 1$, $\Psi(\epsilon v_1) \approx \epsilon\mu_1 v_1 + O(\epsilon^3)$; the symmetry $\omega \mapsto -\omega$ produces $\pm\hat{m}^*$.

Part (iii): Expanding $\hat{m} = \tanh(J\hat{m})$ along v_1 with $\hat{m} = \epsilon v_1$ and using $\tanh(x) = x - x^3/3 + O(x^5)$ gives $\epsilon \approx \mu_1\epsilon - (\mu_1\epsilon)^3/3$, so $\epsilon^* \approx \sqrt{3(\mu_1 - 1)/\mu_1^3}$. □

Corollary 2.8 (Gershgorin sufficient criterion)**.** *A sufficient condition for linear stability of $\hat{m} = 0$ is $\beta_{\text{eff}}(J) < 1$, since $\|J\|_{\text{op}} \leq \beta_{\text{eff}}(J)$. The condition is conservative: it can fail while $\|J\|_{\text{op}} < 1$ still holds.*

Remark 2.9 (Gershgorin slack: explicit non-tightness)**.** The gap between $\beta_{\text{eff}}(J) = d_{\max} J_{\max}$ and $\|J\|_{\text{op}}$ can be arbitrary. Consider $M = 20$ with J_{ij} i.i.d. Rademacher sign times $J_{\max} = 0.1$ on a d-regular graph with $d_{\max} = 20$. Then $\beta_{\text{eff}} = 2.0$, violating Corollary 2.8's sufficient condition, while $\|J\|_{\text{op}}$ is concentrated around $2J_{\max}\sqrt{d_{\max}} \approx 0.89$ (Wigner semicircle for bounded symmetric matrices), so linear stability of $\hat{m} = 0$ is nevertheless preserved. The Gershgorin bound is tight only for matrices with coherent (ferromagnetic) row sums; for signed or random-sign networks it overstates $\|J\|_{\text{op}}$ by a factor $O(\sqrt{d_{\max}})$. Consequently, the estimation results of Sections 3-6 require $\|J\|_{\text{op}} < 1$, which is weaker than the Gershgorin criterion and often materially easier to satisfy.

Remark 2.10 (Signed J)**.** Theorem 2.7(ii) restricts to irreducible non-negative J, and anti-correlated events give $J_{ij} < 0$. The estimation theorems of Sections 3-6 do *not* depend on the bifurcation structure of the signed case; they require only that the pair (i, j) operate in a regime where the two-site Fisher information is positive-definite, which is guaranteed by $\|J\|_{\text{op}} < 1$ (Proposition 2.11) regardless of the sign pattern.

Proposition 2.11 (Linearised susceptibility)**.** *Define $\chi_{ij} = \partial\hat{m}_i/\partial h_j\big|_{h=0}$. In the stable regime $\|J\|_{\text{op}} < 1$,*

$$\chi = (I - J)^{-1}\big|_{\hat{m}=0}, \qquad \|\chi\|_{\text{op}} \leq \frac{1}{1 - \|J\|_{\text{op}}} \leq \frac{1}{1 - \beta_{\text{eff}}(J)}. \tag{8}$$

Within the mean-field approximation, $\|\chi\|_{\text{op}}$ diverges as $\|J\|_{\text{op}} \to 1^-$.

Proof. Differentiating (6) at $\hat{m} = 0$: $\partial\hat{m}_i/\partial h_j = \delta_{ij} + \sum_k J_{ik}\,\partial\hat{m}_k/\partial h_j$, i.e. $\chi = I + J\chi$, so $\chi = (I - J)^{-1}$. The operator-norm bound follows from the Neumann series. □

The practical consequence for the estimators below is that the MLE Fisher-information matrix (Theorem 3.6) remains positive-definite and bounded away from singularity as long as the two-site pair (i, j) operates below its local instability threshold; pairs approaching $\|J\|_{\text{op}} = 1$ (not considered in the present paper) would require a separate critical-regime analysis.

3 Identifiability from Marginal and Parlay Prices

3.1 Closed-form inversion at $M = 2$

A parlay trade on pair (i, j) pays \$1 if $\omega_i = +1$ and $\omega_j = +1$ jointly, zero otherwise; its market price identifies the joint probability

$$\pi_{ij} \;=\; \mathbb{P}(\omega_i = +1, \omega_j = +1).$$

Proposition 3.1 (Parlay inversion: closed form at $M = 2$, leading-order at $M > 2$). *Given marginals p_i, p_j and parlay price π_{ij} in the two-event model,*

$$J_{ij} \;=\; \frac{1}{4} \ln \frac{\pi_{ij}\,(1 - p_i - p_j + \pi_{ij})}{(p_i - \pi_{ij})(p_j - \pi_{ij})}. \tag{9}$$

For $M > 2$, applying (9) to a pair (i, j) gives a consistent estimator whose bias relative to the true J^*_{ij} is $O(J^2_{\max})$; the sharpened form is in Theorem 5.1.

Proof. The four-state Boltzmann distribution at $M = 2$ gives

$$\begin{aligned}
\mathbb{P}(+1, +1) &= Z^{-1} \exp(h_1 + h_2 + J_{12}),\\
\mathbb{P}(+1, -1) &= Z^{-1} \exp(h_1 - h_2 - J_{12}),\\
\mathbb{P}(-1, +1) &= Z^{-1} \exp(-h_1 + h_2 - J_{12}),\\
\mathbb{P}(-1, -1) &= Z^{-1} \exp(-h_1 - h_2 + J_{12}).
\end{aligned}$$

The cross-ratio

$$\frac{\mathbb{P}(+1, +1)\,\mathbb{P}(-1, -1)}{\mathbb{P}(+1, -1)\,\mathbb{P}(-1, +1)} \;=\; \exp(4J_{12})$$

identifies J_{12} up to the four observed probabilities. Substituting $\mathbb{P}(+1, +1) = \pi_{12}$, $\mathbb{P}(+1, -1) = p_1 - \pi_{12}$, $\mathbb{P}(-1, +1) = p_2 - \pi_{12}$, $\mathbb{P}(-1, -1) = 1 - p_1 - p_2 + \pi_{12}$ gives (9). □

3.2 Exact identifiability and the admissible parlay region

The closed form (9) is real-valued only if every factor in the cross-ratio is strictly positive; this constrains π_{ij} relative to (p_i, p_j).

Theorem 3.2 (Exact identifiability of the two-event Ising model). *Fix marginals $p_i, p_j \in (0, 1)$. Define the* admissible parlay region

$$\mathcal{A}(p_i, p_j) \;:=\; \{\pi \in (0, 1) \;:\; \max(0, p_i + p_j - 1) \;<\; \pi \;<\; \min(p_i, p_j)\}. \tag{10}$$

Then:

(i) For every $\pi_{ij} \in \mathcal{A}(p_i, p_j)$, the map $\Phi : (p_i, p_j, \pi_{ij}) \mapsto (h_i, h_j, J_{ij})$ defined by

$$h_i = \tfrac{1}{2}\,\mathrm{logit}(p_i), \qquad h_j = \tfrac{1}{2}\,\mathrm{logit}(p_j), \qquad J_{ij} = \tfrac{1}{4} \ln \frac{\pi_{ij}(1 - p_i - p_j + \pi_{ij})}{(p_i - \pi_{ij})(p_j - \pi_{ij})},$$

is a bijection onto $\mathbb{R}^3$. The two-event Ising parameters (h_i, h_j, J_{ij}) are therefore uniquely determined by the triple (p_i, p_j, π_{ij}) inside the admissible region.

(ii) The admissible region $\mathcal{A}(p_i, p_j)$ is exactly the image under Φ^{-1} of $\mathbb{R}^3$; every triple $(p_i, p_j, \pi_{ij}) \in (0, 1)^2 \times \mathcal{A}(p_i, p_j)$ corresponds to a well-defined two-event Ising distribution, and no triple outside this region does.

(iii) The Fréchet-Hoeffding bounds $\max(0, p_i + p_j - 1) \le \pi \le \min(p_i, p_j)$ *are attained in the limit* $J_{ij} \to \pm\infty$*; the strict inequalities in* $\mathcal{A}$ *correspond to finite* J_{ij}*.*

(iv) $\pi_{ij} = p_i p_j$ *iff* $J_{ij} = 0$ *(independence).* $\pi_{ij} > p_i p_j$ *iff* $J_{ij} > 0$*;* $\pi_{ij} < p_i p_j$ *iff* $J_{ij} < 0$*.*

Proof. (i) *Injectivity.* h_i and h_j are determined by p_i, p_j separately through the logit. Given (p_i, p_j), the map $\pi \mapsto J$ via (9) has derivative

$$\frac{dJ}{d\pi} = \frac{1}{4} \cdot \frac{d}{d\pi} \ln \frac{\pi(1 - p_i - p_j + \pi)}{(p_i - \pi)(p_j - \pi)} = \frac{1}{4}\left[\frac{1}{\pi} + \frac{1}{1 - p_i - p_j + \pi} + \frac{1}{p_i - \pi} + \frac{1}{p_j - \pi}\right],$$

which is strictly positive on $\mathcal{A}$ because each of the four denominators is positive exactly on $\mathcal{A}$. Hence J is a strictly monotone (increasing) function of π on $\mathcal{A}$.

Surjectivity. At the lower boundary $\pi \downarrow \max(0, p_i + p_j - 1)$, one of the denominators $\{\pi, 1 - p_i - p_j + \pi\}$ tends to zero, forcing $J \to -\infty$. At the upper boundary $\pi \uparrow \min(p_i, p_j)$, one of $\{p_i - \pi,\ p_j - \pi\}$ tends to zero, forcing $J \to +\infty$. Continuity and strict monotonicity give Φ maps $\mathcal{A}$ onto all of $\mathbb{R}$ for J, and onto $\mathbb{R}^2$ for (h_i, h_j). The combined map is therefore a bijection $\mathcal{A}_{(p_i,p_j)} \times (0,1)^2 \to \mathbb{R}^3$.

(ii) The admissible region is the image of $\mathbb{R}^3$ under the forward map $(h_i, h_j, J_{ij}) \mapsto (p_i, p_j, \pi_{ij})$. Compute the forward map: for any $(h_i, h_j, J_{ij}) \in \mathbb{R}^3$, the two-site Boltzmann normalisation yields strictly positive values for all four joint probabilities, hence marginals $p_i, p_j \in (0, 1)$ and a parlay probability π_{ij} lying strictly between the Fréchet bounds. Conversely, a triple outside the Fréchet bounds cannot correspond to a valid joint distribution at all.

(iii) Direct computation at the boundaries.

(iv) At $J_{ij} = 0$, the Boltzmann factors are $\exp(\pm h_i \pm h_j)$ and the distribution factorises to $\mathbb{P}(\omega_i, \omega_j) = \mathbb{P}(\omega_i)\mathbb{P}(\omega_j)$, so $\pi_{ij} = p_i p_j$. Strict monotonicity of $J(\pi)$ from (i) gives the signed equivalences. □

Remark 3.3 (Fréchet-Hoeffding bounds)**.** The admissible region $\mathcal{A}(p_i, p_j)$ is exactly the interior of the Fréchet-Hoeffding compatibility region for two Bernoulli margins: the strict inequalities arise because the Ising parameterisation has no boundary atoms. A market that quotes (p_i, p_j, π_{ij}) outside the closed Fréchet region is inconsistent with every joint distribution on $\{-1, +1\}^2$. A market at the boundary is consistent and has infinite $|J_{ij}|$; the Ising form is not designed for the boundary case and a separate limiting analysis is required.

Example 3.4 (Numerical illustration)**.** Suppose $p_1 = 0.65$, $p_2 = 0.40$. The admissible parlay region is $\mathcal{A} = (0.05, 0.40)$.

- Independence: $\pi_{12} = p_1 p_2 = 0.26$, giving $J_{12} = 0$ from (9).
- Positive coupling: $\pi_{12} = 0.31$ yields $J_{12} = \frac{1}{4}\ln[0.31 \cdot 0.06/(0.34 \cdot 0.09)] \approx +0.12$.
- Negative coupling: $\pi_{12} = 0.21$ yields $J_{12} = \frac{1}{4}\ln[0.21 \cdot (-0.04)/(0.44 \cdot 0.19)]$, but $1 - p_1 - p_2 + \pi = -0.04 < 0$ lies outside $\mathcal{A}$, so this triple is infeasible. The lower boundary of $\mathcal{A}$ is $\max(0, p_1 + p_2 - 1) = 0.05$; any $\pi < 0.05$ violates Fréchet. A feasible negative-coupling example is $\pi_{12} = 0.18$: $J_{12} = \frac{1}{4}\ln[0.18 \cdot 0.13/(0.47 \cdot 0.22)] \approx -0.17$.

3.3 Maximum-likelihood estimation from parlay trades

The closed-form inversion (9) takes π_{ij} as known; in practice π_{ij} is estimated from a finite stream of parlay outcomes. The MLE on the empirical joint-outcome counts formalises the finite-sample estimator. The resulting model is the classical 2×2 log-linear exponential family (Bishop-Fienberg-Holland [34], Agresti [35], §9.2) with sufficient statistics

$(\omega_i, \omega_j, \omega_i\omega_j)$; the contribution of this section is the explicit Schur-complement form for the efficient variance of $\hat{J}_{ij}$ in the off-diagonal regime.

Definition 3.5 (Parlay observation model)**.** Let N_p parlay contracts on pair (i, j) be resolved over a window. Each resolution reports the joint outcome $(\omega_i, \omega_j) \in \{-1, +1\}^2$; the resolutions are i.i.d. draws from the pair-marginal of the full M-site Ising distribution, conditional on a parlay contract having been listed on the pair (see Assumption 7.1). Let $N_{++}, N_{+-}, N_{-+}, N_{-}$ count the four joint outcomes with $N_p = \sum N_{\cdot\cdot}$. Under the two-site Ising model with parameter $\theta = (h_i, h_j, J_{ij})$, the log-likelihood is

$$\ell(\theta) \;=\; \sum_{s_i, s_j \in \{\pm 1\}} N_{s_i s_j}\big(h_i s_i + h_j s_j + J_{ij} s_i s_j\big) \;-\; N_p \ln Z(\theta), \tag{11}$$

with $Z(\theta) = \sum_{s_i, s_j} \exp(h_i s_i + h_j s_j + J_{ij} s_i s_j)$.

Theorem 3.6 (MLE consistency, asymptotic normality, and efficient variance)**.** *Let $\hat{\theta}_{N_p} = \arg\max_\theta \ell(\theta)$ with truth θ^* in a compact interior of the parameter space, and N_p i.i.d. parlay trades with finite $|J^*_{ij}|$. Then:*

(i) (Consistency) $\hat{\theta}_{N_p} \to \theta^*$ in probability as $N_p \to \infty$.

(ii) (Asymptotic normality) $\sqrt{N_p}(\hat{\theta}_{N_p} - \theta^*) \xrightarrow{d} \mathcal{N}(0, I(\theta^*)^{-1})$, where $I(\theta^*) = \mathrm{Cov}_{\theta^*}(T)$ with sufficient statistics $T = (\omega_i, \omega_j, \omega_i\omega_j)$:

$$I(\theta^*) \;=\; \begin{pmatrix} 1 - m_i^2 & C_{ij} - m_i m_j & m_j - m_i C_{ij} \\ C_{ij} - m_i m_j & 1 - m_j^2 & m_i - m_j C_{ij} \\ m_j - m_i C_{ij} & m_i - m_j C_{ij} & 1 - C_{ij}^2 \end{pmatrix}, \tag{12}$$

with $m_i = \mathbb{E}[\omega_i]$, $m_j = \mathbb{E}[\omega_j]$, $C_{ij} = \mathbb{E}[\omega_i\omega_j]$ under $\mathbb{P}_{\theta^*}$.

(iii) (Efficient variance: Schur complement form) Partition $I(\theta^*) = \begin{pmatrix} A & b \\ b^\top & I_{33} \end{pmatrix}$ with $A \in \mathbb{R}^{2\times 2}$ the field-sector block, $b = (m_j - m_i C_{ij},\ m_i - m_j C_{ij})^\top$, and $I_{33} = 1 - C^{*2}_{ij}$. The efficient asymptotic variance of $\hat{J}_{ij}$ is

$$\mathrm{Var}_{\mathrm{eff}}(\hat{J}_{ij}) \;=\; \frac{1}{N_p}\,[I(\theta^*)^{-1}]_{33} \;=\; \frac{1}{N_p\,(I_{33} - b^\top A^{-1} b)} \;\geq\; \frac{1}{N_p\,(1 - C^{*2}_{ij})}, \tag{13}$$

with equality iff $b = 0$. The marginal bound $1/[N_p(1 - C^{*2}_{ij})]$ ignores the correlation between the J-direction and the field sector; the Schur correction $b^\top A^{-1} b$ is a constant-factor reduction of the denominator, not *an $O(1/N_p^2)$ effect, and is generically non-zero when $m_i m_j \neq 0$.*

Proof. The two-site Ising model is a regular exponential family with sufficient statistics $T = (\omega_i, \omega_j, \omega_i\omega_j)$ and natural parameters $\theta = (h_i, h_j, J_{ij})$. Consistency (i) follows from standard exponential-family MLE theory: the log-likelihood is strictly concave in θ (Hessian $-N_p I(\theta) \prec 0$) and the score function has mean zero at the truth (Lehmann-Casella [23], Ch. 6); a detailed proof via Wald's theorem is given as Theorem 7.2 below.

For (ii), the Fisher information of an exponential family equals $\mathrm{Cov}_{\theta^*}(T)$. Compute entry-wise using $\omega_k^2 \equiv 1$:

$$I_{11} = \mathrm{Var}(\omega_i) = 1 - m_i^2, \qquad I_{12} = \mathrm{Cov}(\omega_i, \omega_j) = C_{ij} - m_i m_j,$$
$$I_{22} = 1 - m_j^2, \qquad I_{33} = \mathrm{Var}(\omega_i\omega_j) = 1 - C_{ij}^2,$$
$$I_{13} = \mathrm{Cov}(\omega_i, \omega_i\omega_j) = \mathbb{E}[\omega_i^2\omega_j] - m_i C_{ij} = m_j - m_i C_{ij}, \qquad I_{23} = m_i - m_j C_{ij}.$$

For (iii), the standard 3×3 block inverse formula gives $[I^{-1}]_{33} = (I_{33} - b^\top A^{-1} b)^{-1}$; positive semidefiniteness of A^{-1} gives $b^\top A^{-1} b \geq 0$, with equality iff $b = 0$. Scaling by $1/N_p$ gives (13). $\square$

Remark 3.7 (Closed-form estimator vs. MLE)**.** Two estimators are available at $M = 2$ with parlay data: the closed-form estimator (9) (method-of-moments on empirical joint-outcome frequencies) and the MLE of Theorem 3.6. The MLE attains the efficient Cramér-Rao variance (13); the closed-form estimator is consistent and inefficient relative to the MLE, with asymptotic variance exceeding the MLE's by a pair-dependent factor computable from (p_i, p_j, π_{ij}). We prefer the MLE whenever numerical optimisation is available; the closed form remains useful as an initial value and as a sanity-check for the MLE optimiser.

4 Non-Asymptotic Concentration for $\hat{J}_{ij}$

Theorem 3.6(ii) is an asymptotic statement. For decision-making with a finite number of parlay observations, a non-asymptotic concentration bound is needed. The two-site Ising model's sufficient statistics are bounded in $\{-1, +1\}$, so Hoeffding-type concentration applies directly.

Theorem 4.1 (Non-asymptotic concentration of $\hat{J}_{ij}$, local to θ^*)**.** *Under the hypotheses of Theorem 3.6, there exists a constant $C(\theta^*) > 0$, computable from the moments (m_i^*, m_j^*, C_{ij}^*) under $\mathbb{P}_{\theta^*}$, and a radius $\epsilon_0(\theta^*) > 0$ such that for every $\epsilon \in (0, \epsilon_0(\theta^*))$ and $\delta \in (0, 1)$,*

$$\mathbb{P}\big(|\hat{J}_{ij,N_p} - J_{ij}^*| > \epsilon\big) \;\leq\; \delta \quad \textit{whenever} \quad N_p \;\geq\; \frac{C(\theta^*)}{\epsilon^2} \log\frac{2}{\delta}. \tag{14}$$

The constant is

$$C(\theta^*) \;=\; \frac{8}{\big(I_{33} - b^\top A^{-1} b\big)^2 \lambda_{\min}(A)}, \tag{15}$$

with A, b, I_{33} as in Theorem 3.6(iii), and $\lambda_{\min}(A) > 0$ the smallest eigenvalue of the field-sector Fisher block. The radius $\epsilon_0(\theta^*)$ is the largest $r > 0$ such that $I(\theta) \succeq \frac{1}{2} I(\theta^*)$ for all $\|\theta - \theta^*\| \leq r$; by continuity of I, $\epsilon_0(\theta^*)$ is controlled by $1/(1 - \|J^*\|_{\mathrm{op}})$ on the two-site pair. Outside this ball the implicit-function linearisation fails and (14) is no longer asserted. The rate $\epsilon^{-2}\log(1/\delta)$ is the parametric MLE rate for a fixed regular exponential family and is tight up to constants [24].

Proof. Step 1: Hoeffding on the sufficient-statistic averages. Each of $\bar{T}_k := N_p^{-1} \sum_r T_k(\omega^{(r)})$ for $k = 1, 2, 3$ is a bounded average (each $T_k(\omega) \in \{-1, +1\}$), so Hoeffding's inequality gives

$$\mathbb{P}\big(|\bar{T}_k - \mathbb{E}[T_k]| > t\big) \;\leq\; 2\exp(-N_p t^2/2), \qquad k = 1, 2, 3.$$

Union-bounding over the three components and rearranging, with probability $\geq 1 - \delta/2$,

$$\|\bar{T} - \mathbb{E}[T]\|_\infty \;\leq\; \sqrt{\frac{2}{N_p} \log\frac{12}{\delta}}.$$

Step 2: mean-parameter map to natural parameter. The map from mean parameters $\mu := \mathbb{E}[T]$ to natural parameters θ is smooth on the interior of the mean-parameter simplex, with Jacobian $(\partial\theta/\partial\mu) = I(\theta)^{-1}$ (exponential-family duality [24]). Writing $S := I_{33} - b^\top A^{-1} b > 0$ for the Schur complement and applying the standard block-inverse formula [26, §0.8],

$$[I^{-1}]_{33} \;=\; S^{-1}, \qquad [I^{-1}]_{j,3} \;=\; -S^{-1}\,(A^{-1}b)_j \text{ for } j \in \{1, 2\}.$$

Hence the $(J_{ij}, \cdot)$ row of I^{-1} has squared Euclidean norm

$$\|[I^{-1}]_{3,\cdot}\|_2^2 \;=\; S^{-2} + S^{-2}\|A^{-1}b\|_2^2 \;=\; S^{-2}\big(1 + \|A^{-1}b\|_2^2\big).$$

Using $\|b\|_2 \le \sqrt{2}$ (each entry of b is a difference of products of numbers in $[-1,1]$) and the operator-norm bound $\|A^{-1}b\|_2 \le \|b\|_2/\lambda_{\min}(A) \le \sqrt{2}/\lambda_{\min}(A)$,

$$\|[I^{-1}]_{3,\cdot}\|_2^2 \;\le\; S^{-2}\big(1 + 2/\lambda_{\min}(A)^2\big) \;\le\; \frac{3}{S^2\,\lambda_{\min}(A)^2}$$

on the interior, where the final bound uses $\lambda_{\min}(A) \le 1$ (since $A \preceq I_2$ as $A = \mathrm{diag}(1 - m_i^{*2},\, 1 - m_j^{*2}) + (C_{ij}^* - m_i^* m_j^*)(E_{12} + E_{21})$ with each diagonal entry ≤ 1). Taking square roots,

$$\|[I^{-1}]_{3,\cdot}\|_2 \;\le\; \frac{\sqrt{3}}{S\,\lambda_{\min}(A)}.$$

Step 3: propagate the deviation. The MLE solves $\bar{T} = \mathbb{E}_\theta[T]$. By the implicit-function theorem applied to the exponential-family mean-parameter map at θ^*, on the event of Step 1 (within the neighbourhood $\epsilon_0(\theta^*)$ where the Fisher information is uniformly bounded below),

$$|\hat{J}_{ij,N_p} - J_{ij}^*| \;\le\; \|[I^{-1}]_{3,\cdot}\|_2 \cdot \|\bar{T} - \mathbb{E}[T]\|_2 \;\le\; \sqrt{3}\,\|[I^{-1}]_{3,\cdot}\|_2 \cdot \|\bar{T} - \mathbb{E}[T]\|_\infty.$$

Combining with the Hoeffding bound of Step 1 and demanding the right-hand side $\le \epsilon$ yields

$$N_p \;\ge\; \frac{6\,\|[I^{-1}]_{3,\cdot}\|_2^2}{\epsilon^2}\log\frac{12}{\delta}.$$

Absorbing the numerical constants and using the Step 2 bound on $\|[I^{-1}]_{3,\cdot}\|_2$ gives (15). The $\log(2/\delta)$ form in (14) replaces the $\log(12/\delta)$ via absorbing the constant 6 into $C(\theta^*)$.

Step 4: restriction to $\epsilon < \epsilon_0(\theta^*)$. The Fisher matrix $I(\theta)$ is positive-definite on the interior of the parameter space and depends continuously on θ. In the neighbourhood $\{\theta : \|\theta - \theta^*\| \le \epsilon_0\}$ with ϵ_0 chosen so that $I(\theta)$ has smallest eigenvalue at least $\lambda_{\min}(A(\theta^*))/2$, Step 3's implicit-function argument applies; outside that neighbourhood the local linearisation of the mean-parameter map fails. The size $\epsilon_0(\theta^*)$ is controlled by $1/(1 - \|J^*\|_{\mathrm{op}})$ through Proposition 2.11. $\square$

Remark 4.2 (Comparison with the asymptotic variance)**.** The asymptotic variance from Theorem 3.6(iii) is $[(I_{33} - b^\top A^{-1}b)N_p]^{-1}$. Theorem 4.1's finite-sample requirement $N_p \ge C(\theta^*)\epsilon^{-2}\log(2/\delta)$ with $C(\theta^*) = 8/[(I_{33} - b^\top A^{-1}b)^2\lambda_{\min}(A)]$ matches the asymptotic variance up to the $\lambda_{\min}(A)^{-1}$ factor (from the field-sector coupling) and a logarithmic factor in $1/\delta$. In the zero-marginal-bias locus $m_i = m_j = 0$, $b = 0$ and A is $\mathrm{diag}(1 - m_i^2, 1 - m_j^2)$, so $C(\theta^*)$ reduces to the classical $8/(1 - C_{ij}^{*2})^2 \cdot (1 - \max(m_i^2, m_j^2))^{-1}$.

Example 4.3 (Sample complexity for $\epsilon = 0.01$)**.** For a representative regime with $(m_i^*, m_j^*, C_{ij}^*) = (0.3, 0.3, 0.2)$, using the Schur-correction numerics of Remark 7.6 below: $I_{33} - b^\top A^{-1}b \approx 0.847$, $\lambda_{\min}(A) \approx 0.8$, giving $C(\theta^*) \approx 8/(0.847^2 \cdot 0.8) \approx 13.9$. Therefore $N_p \ge 13.9 \cdot \epsilon^{-2} \cdot \log(2/\delta)$. At $\epsilon = 0.01$, $\delta = 0.01$: $13.9 \cdot 10^4 \cdot \log 200 \approx 1.39 \times 10^5 \cdot 5.30 \approx 7.4 \times 10^5$ parlays. At looser targets $\epsilon = 0.05$, $\delta = 0.05$: $13.9 \cdot 400 \cdot \log 40 \approx 5560 \cdot 3.69 \approx 2.1 \times 10^4$ parlays. These are sufficient-sample bounds from the Hoeffding route; the MLE asymptotic variance $[(I_{33} - b^\top A^{-1}b)N_p]^{-1}$ is tighter by a factor of ~ 8 in the lead constant, so the actual sample requirement to achieve the stated accuracy in expectation is closer to $N_p \approx 10^5$ (for $\epsilon = 0.01$), with the 7.4×10^5 figure providing the high-probability upper bound.

5 Sharpened $M > 2$ Embedding Bias

When the two-site estimator of Proposition 3.1 is applied to a pair (i, j) inside a network of $M > 2$ coupled events, the resulting estimator is biased: the estimator assumes the two-site Boltzmann distribution, but the true marginal distribution of (ω_i, ω_j) is obtained by marginalising the full M-site distribution. The leading-order bias is computable.

Theorem 5.1 (Two-site embedding bias: sharpened form)**.** *Let $J^* \in \mathbb{R}^{M\times M}$ be the true symmetric interaction matrix with $J_{\max} = \max_{k,l} |J^*_{kl}|$ and mean degree $D = \max_i |\{k : J^*_{ik} \neq 0\}|$; let (m^*_k) be the marginals under the full M-site Boltzmann distribution. Suppose the closed-form estimator (9) is applied to the pair (i, j) with $\pi_{ij} = \mathbb{P}_{J^*}(\omega_i = +1, \omega_j = +1)$ (i.e. the true M-site joint). Let $\hat{J}^{(M>2)}_{ij}$ denote the resulting closed-form estimate. Then*

$$\hat{J}^{(M>2)}_{ij} \;=\; J^*_{ij} \;+\; \Delta_{ij}, \qquad \Delta_{ij} \;=\; -\frac{1}{2}\sum_{k\neq i,j} J^*_{ik}J^*_{jk}\,(1-m^{*2}_k) \;+\; R_{ij}, \tag{16}$$

where, under the Plefka / Thouless-Anderson-Palmer expansion [32, 33], the remainder satisfies

$$|R_{ij}| \;\leq\; 2\,D\,J^3_{\max} \qquad \textit{whenever} \qquad D J_{\max} \leq 1/2. \tag{17}$$

Constants and proof strategy are stated in Remark 5.3.

Proof. Write the full M-site Hamiltonian as $H = H_{ij} + H_{\text{other}} + H_{\text{coupling}}$ where

$$\begin{aligned} H_{ij}(\omega_i, \omega_j) &= -h_i\omega_i - h_j\omega_j - J^*_{ij}\omega_i\omega_j, \\ H_{\text{other}}(\omega_{-ij}) &= -\sum_{k\neq i,j} h_k\omega_k - \sum_{\{k,l\}\cap\{i,j\}=\emptyset} J^*_{kl}\omega_k\omega_l, \\ H_{\text{coupling}}(\omega) &= -\sum_{k\neq i,j}(J^*_{ik}\omega_k)\omega_i - \sum_{k\neq i,j}(J^*_{jk}\omega_k)\omega_j. \end{aligned}$$

Marginalising over ω_{-ij}, the reduced two-site distribution is

$$\mathbb{P}(\omega_i, \omega_j) \;\propto\; \exp\!\big(h_i\omega_i + h_j\omega_j + J^*_{ij}\omega_i\omega_j\big)\cdot \mathbb{E}_{\omega_{-ij}}\Big[\exp\sum_{k\neq i,j}(J^*_{ik}\omega_i + J^*_{jk}\omega_j)\omega_k\Big].$$

The inner expectation is a partition function over ω_{-ij} with effective fields $\tilde{h}_k = h_k + J^*_{ik}\omega_i + J^*_{jk}\omega_j$ (plus the other-other couplings). Cumulant expansion of the log of this partition function around $\omega_i = \omega_j = 0$ gives

$$\ln \mathbb{E}_{\omega_{-ij}}[\ldots] \;=\; \ln Z_{\text{other}} \;+\; \sum_{k\neq i,j} m^*_k(J^*_{ik}\omega_i + J^*_{jk}\omega_j) \;+\; \tfrac{1}{2}(J^*_{ik}\omega_i + J^*_{jk}\omega_j)^2(1-m^{*2}_k) \;+\; O(J^3_{\max}).$$

The first-order term renormalises (h_i, h_j) to $\tilde{h}_i = h_i + \sum_k J^*_{ik}m^*_k$, $\tilde{h}_j$ analogously. The second-order term expands as

$$\tfrac{1}{2}\sum_k\Big[J^{*2}_{ik}(1-m^{*2}_k) + 2J^*_{ik}J^*_{jk}(1-m^{*2}_k)\omega_i\omega_j + J^{*2}_{jk}(1-m^{*2}_k)\Big].$$

The cross-term $2J^*_{ik}J^*_{jk}(1-m^{*2}_k)\omega_i\omega_j$ contributes a correction to the effective two-site interaction. Comparing to the two-site form, $\omega_i\omega_j$ has effective coupling $J^*_{ij} + \sum_k J^*_{ik}J^*_{jk}(1-m^{*2}_k)$ when measured by the two-site estimator's log-linear inversion. Negating (because of the sign of the cumulant relative to $-H$ in (3)) gives

$$\Delta_{ij} \;=\; -\tfrac{1}{2}\sum_{k\neq i,j} J^*_{ik}J^*_{jk}(1-m^{*2}_k) \;+\; R_{ij},$$

matching (16).

The remainder R_{ij} collects third- and higher-order Plefka cumulants of ω_{-ij} weighted by products of J^* couplings incident on (i,j). The third-order Plefka term [32, 33] contributes $\sum_k J^*_{ik}J^*_{jk}\gamma_k \cdot \max_\ell |J^*_{k\ell}|$ with $|\gamma_k| \le (1-m_k^{*2})^2 \le 1$; only indices k in the support of $J^*_{i\cdot}$ and $J^*_{j\cdot}$ contribute, and the degree bound $|\{k : J^*_{ik} \ne 0\}| \le D$ gives $|R_{ij}| \le 2DJ^3_{\max}$. Higher Plefka orders are controlled by the same degree-and-operator-norm expansion under the weak-coupling condition $DJ_{\max} \le 1/2$, which places the iteration inside its radius of convergence. □

Corollary 5.2 (Bias is negligible in the sparse small-coupling regime)**.** *Suppose $|\{k : J^*_{ik} \ne 0\}| \cdot |\{k : J^*_{jk} \ne 0\}| \le D^2$ and $DJ_{\max} \le 1/2$. Then $|\Delta_{ij}| \le \frac{1}{2}D^2J^2_{\max} + 2DJ^3_{\max} \le (DJ_{\max})^2$, which is a fraction $D^2J_{\max}/|J^*_{ij}|$ of the true coupling. At $D = 5$ and $J_{\max} = 0.1$: $|\Delta_{ij}| \le 0.125 \cdot J_{\max}$ in the worst case; negligible for typical $|J^*_{ij}|$ at or above $J_{\max}/2$.*

Remark 5.3 (Remainder via TAP expansion)**.** The remainder R_{ij} in Theorem 5.1 collects third- and higher-order cumulants of ω_{-ij} weighted by products of J^* couplings. The Plefka / Thouless-Anderson-Palmer (TAP) expansion of the free energy of a pairwise-Ising model [32, 13, 33] organises these corrections systematically: the first Plefka order recovers the mean-field coupling renormalisation $\tilde{h}_i = h_i + \sum_k J^*_{ik}m^*_k$, the second order yields the quadratic cross-term $-\frac{1}{2}\sum_k J^*_{ik}J^*_{jk}(1-m_k^{*2})$ of (16), and the third order yields the Onsager reaction correction with coefficient depending on the local curvature $(1-m_k^{*2})^2$ of the single-site entropy. In the weak-coupling regime $J_{\max} \le 1/(2(M-2))$ the third-order term is bounded in ℓ^1 by $\sum_k |J^*_{ik}J^*_{jk}|\ \max_\ell |J^*_{k\ell}|\,(1-m_k^{*2})^2 \le J^2_{\max}\sum_k |J^*_{ik}| \cdot \max_\ell |J^*_{k\ell}|$, giving $|R_{ij}| \le 2J^3_{\max}\sum_{k\ne i,j}|\{l : J^*_{kl} \ne 0\}|$. For a sparse network with mean degree D this is $|R_{ij}| \le 2DJ^3_{\max}$, an $O(D)$ bound rather than the loose $O((M-2))$ bound of the direct cumulant count. The Plefka route also makes explicit that the Onsager correction vanishes when two of the three indices coincide, which the generic-constant bound of (17) does not exploit.

6 Robustness to Quoted-Marginal Noise

The identification $h_i = \frac{1}{2}\operatorname{logit}(p_i)$ relies on the market price p_i coinciding with the Boltzmann marginal $\mathbb{P}(\omega_i = +1)$. Under a market-scoring-rule market [2] with liquidity parameter b, this identification holds at the equilibrium price where traders have no profitable trades. Away from equilibrium, under inventory imbalance, liquidity shocks, or transient order-book imbalances, the quoted price deviates from the Boltzmann marginal. The following result bounds how those deviations propagate into the inferred (h_i, J_{ij}).

Theorem 6.1 (Robustness of parlay inversion)**.** *Let (p_i, p_j, π_{ij}) be the true triple and $(\tilde{p}_i, \tilde{p}_j, \tilde{\pi}_{ij})$ the quoted triple with $|\tilde{p}_i - p_i| \le \eta$, $|\tilde{p}_j - p_j| \le \eta$, $|\tilde{\pi}_{ij} - \pi_{ij}| \le \eta$. Suppose the true triple lies in the κ-interior of the admissible region $\mathcal{A}(p_i, p_j)$, i.e. each of $\{\tilde{\pi}_{ij},\ \tilde{p}_i - \tilde{\pi}_{ij},\ \tilde{p}_j - \tilde{\pi}_{ij},\ 1 - \tilde{p}_i - \tilde{p}_j + \tilde{\pi}_{ij}\}$ exceeds $\kappa > 0$ for all quoted triples within the noise ball. Then the inferred parameters under the closed form (9) satisfy*

$$|\tilde{h}_i - h_i| \le \frac{\eta}{2\min(p_i, 1-p_i)\cdot\max(1-\eta/\min(p_i,1-p_i), 1/2)} = O(\eta), \qquad (18)$$

$$|\tilde{J}_{ij} - J_{ij}| \le \frac{\eta}{\kappa}. \qquad (19)$$

In the well-quoted regime $\eta \ll \kappa$ and $\eta \ll \min(p_i, 1-p_i, p_j, 1-p_j)$, the inferred parameters inherit the $O(\eta)$ precision of the quoted prices, with the Lipschitz constant of (19) controlled by the distance to the admissible-region boundary.

Proof. For h: $h_i = \frac{1}{2}\operatorname{logit}(p_i)$ has derivative $dh_i/dp_i = 1/(2p_i(1-p_i))$, bounded on any compact subinterval $[p_i - \eta, p_i + \eta] \subset (0,1)$ by $1/[2\min(p_i, 1-p_i)(1-\eta/\min)]$. A first-order Taylor expansion gives (18).

For J: from Proposition 3.1,

$$J_{ij}(p_i, p_j, \pi) \;=\; \tfrac{1}{4}\big[\ln \pi + \ln(1 - p_i - p_j + \pi) - \ln(p_i - \pi) - \ln(p_j - \pi)\big].$$

The partial derivatives at (p_i, p_j, π) satisfy

$$\left|\frac{\partial J_{ij}}{\partial \pi}\right| \;\le\; \frac{1}{4}\left[\frac{1}{\pi} + \frac{1}{1 - p_i - p_j + \pi} + \frac{1}{p_i - \pi} + \frac{1}{p_j - \pi}\right] \;\le\; \frac{1}{\kappa},$$

using the κ-interior assumption. The mixed partials in p_i, p_j are similarly bounded. Lipschitz-propagating the η-ball of the quoted triple through this Jacobian gives (19). □

Remark 6.2 (Boundary regime)**.** The robustness bound degrades at the admissible-region boundary (Theorem 3.2(iii)): when π is close to $\max(0, p_i + p_j - 1)$ or $\min(p_i, p_j)$, the Lipschitz constant $1/\kappa$ diverges, reflecting that $|J_{ij}|$ is going to infinity. Markets quoting boundary parlay prices are informative about the *sign* of J_{ij} but provide poor magnitude estimates; in practice a quoted-price "de-boundarisation" procedure (e.g., shrinking the quoted triple toward the independence point by a factor of $1-\alpha$ for a small α) can stabilise the estimator at the cost of a controlled bias. This is a market-microstructure design choice not pursued further here.

7 Fortification: Consistency and Efficient Cramér-Rao Bound

This section provides a self-contained proof apparatus for Theorem 3.6: (1) Wald-type consistency via uniform convergence of the log-likelihood, (2) the Schur complement identity for the Fisher inverse, and (3) numerical illustration of the Schur correction in the off-diagonal regime $m_i m_j \neq 0$.

7.1 Notation and regularity

Fix a pair (i,j) and suppress indices: $\omega = (\omega_i, \omega_j)$, $\theta = (h_i, h_j, J_{ij}) \in \Theta \subset \mathbb{R}^3$. Let $\mathbb{P}_\theta$ denote the two-site Ising distribution with density

$$p_\theta(\omega) \;=\; \exp\!\big(h_i\omega_i + h_j\omega_j + J_{ij}\omega_i\omega_j - A(\theta)\big), \qquad A(\theta) = \ln\sum_{\omega} \exp(\cdot).$$

The sufficient statistics are $T(\omega) = (\omega_i, \omega_j, \omega_i\omega_j) \in \{-1,+1\}^3$, and the model is a minimal regular exponential family on the 4-point support $\{-1,+1\}^2$ with three-dimensional natural parameter space $\Theta = \mathbb{R}^3$ (full-rank, since the three sufficient statistics are affinely independent on the support) [24, Ch. 8][25, Ch. 2]. Minimality, together with $\mathbb{R}^3$ being the full domain of finiteness of the log-partition function A, places the model in the canonical regular-exponential-family setting of [24, Thm. 3.6]. The mean-parameter map $\theta \mapsto \mu(\theta) := \mathbb{E}_\theta T$ is consequently a C^∞ diffeomorphism from $\mathbb{R}^3$ onto the interior of the convex hull of $T(\{-1,+1\}^2) \subset \mathbb{R}^3$, with Jacobian $d\mu/d\theta = I(\theta) = \operatorname{Cov}_\theta(T) \succ 0$ everywhere in Θ. Injectivity of $\theta \mapsto \mathbb{P}_\theta$ on Θ follows immediately.

Assumption 7.1 (Regularity)**.** (R1) θ^* lies in a compact convex subset $\Theta^* \subset \Theta$ with non-empty interior. (R2) Conditional on a parlay contract existing on pair (i,j), the stream of resolved joint outcomes $\omega^{(1)}, \ldots, \omega^{(N_p)}$ is i.i.d. from $\mathbb{P}_{\theta^*}$ restricted to $(\omega_i, \omega_j) \in \{-1,+1\}^2$; in particular, the listing/delisting process of parlay contracts is not modelled here, and the estimator output $\hat{\theta}$ is not fed back into the market over the observation window (batch regime). (R3) θ^* is identifiable: for all $\theta' \in \Theta^* \setminus \{\theta^*\}$, $\mathbb{P}_{\theta'} \neq \mathbb{P}_{\theta^*}$. (R4) Causal sufficiency: the observed events $\{e_1, \ldots, e_M\}$ are causally sufficient for the joint distribution on $\{-1, +1\}^M$, i.e. no latent variable affects more than one observed event's outcome. (R4) is not implied by (R1)-(R3); relaxing it requires auxiliary assumptions that separate direct interaction from common-confounder dependence (see [36] for the identifiability theory under latents).

Identifiability follows from strict convexity of the log-partition function A on the interior of Θ [24, Thm. 1.13]: distinct natural parameters give distinct mean parameters $\mathbb{E}_\theta T$.

7.2 Consistency via Wald's theorem

Theorem 7.2 (Strong consistency of the parlay MLE). *Under Assumption 7.1, the MLE $\hat{\theta}_{N_p} = \arg\max_\theta \ell_{N_p}(\theta)$ from (11) is strongly consistent: $\hat{\theta}_{N_p} \xrightarrow{a.s.} \theta^*$ as $N_p \to \infty$.*

Proof. We verify the hypotheses of the Newey-McFadden modernisation of Wald's theorem [21, Thm. 2.5]:

(W1) (Compactness) Θ^* is compact by (R1).
(W2) (Continuity) $\theta \mapsto \ln p_\theta(\omega)$ is continuous on Θ^* uniformly in ω (finite outcome space and continuity of A).
(W3) (Uniform LLN) $\sup_{\theta \in \Theta^*} |N_p^{-1} \ell_{N_p}(\theta) - \mathbb{E}_{\theta^*} \ln p_\theta(\omega)| \xrightarrow{a.s.} 0$ by Jennrich [22], Theorem 2 (uniform SLLN for compactly-indexed i.i.d. sums with continuous integrand; dominated-convergence envelope is trivial over $\{-1, +1\}^2$).
(W4) (Unique population maximiser) $M(\theta) := \mathbb{E}_{\theta^*} \ln p_\theta(\omega) = -\mathrm{KL}(\mathbb{P}_{\theta^*} \| \mathbb{P}_\theta) + \mathbb{E}_{\theta^*} \ln p_{\theta^*}(\omega)$ is uniquely maximised at θ^* by Gibbs' inequality, using (R3).

Wald's theorem gives $\hat{\theta}_{N_p} \xrightarrow{a.s.} \theta^*$. □

Remark 7.3 (Exponential-family shortcut). Because the model is an exponential family, an alternative route uses the strict concavity of ℓ_{N_p} [23, Ch. 2]: the Hessian is $-N_p I(\theta) \prec 0$ everywhere in the interior, so the MLE is unique and solves the moment equation $\bar{T} = \mathbb{E}_\theta T$. The SLLN gives $\bar{T} \to \mathbb{E}_{\theta^*} T$, and continuity of the mean-parameter inverse gives $\hat{\theta} \to \theta^*$. The Wald route is presented because it extends to misspecified settings (the KL projection) and to models with nuisance parameters, which matter when the two-site estimator is embedded in an M-site network (Section 5).

7.3 Schur complement and efficient Cramér-Rao

Lemma 7.4 (Schur complement). *Let $I = \begin{pmatrix} A & b \\ b^\top & d \end{pmatrix} \in \mathbb{R}^{3\times 3}$ be positive-definite with $A \in \mathbb{R}^{2\times 2}$ and $d > 0$. Then: (i) the Schur complement $S := d - b^\top A^{-1} b > 0$; (ii) $[I^{-1}]_{33} = S^{-1}$; (iii) $S \leq d$ with equality iff $b = 0$.*

Proof. (i) Standard property of the Schur complement of a principal block of a positive-definite matrix [26, Thm. 7.7.6]. (ii) Block-inverse formula [26, §0.8]. (iii) $A^{-1} \succ 0$ gives $b^\top A^{-1} b \geq 0$ with equality iff $b = 0$. □

Theorem 7.5 (Efficient Cramér-Rao bound for $\hat{J}_{ij}$). *Under Assumption 7.1, with Fisher information $I(\theta^*) = \mathrm{Cov}_{\theta^*}(T)$ partitioned as $\begin{pmatrix} A & b \\ b^\top & I_{33} \end{pmatrix}$:*

(a) $\mathrm{Var}_\infty(\hat{J}_{ij}) = [N_p(I_{33} - b^\top A^{-1} b)]^{-1} \geq [N_p(1 - C_{ij}^{*2})]^{-1}$, *with equality to the right-hand bound iff $b = 0$.*
(b) If $m_i^* m_j^* \neq 0$, generically $b \neq 0$; the Schur correction is strictly positive.

(c) *$b = 0$ iff $m_j^* = m_i^* C_{ij}^*$ and $m_i^* = m_j^* C_{ij}^*$ simultaneously, which holds at the zero-mean locus $m_i^* = m_j^* = 0$ or the degenerate locus $|C_{ij}^*| = 1$ (boundary of valid Ising distributions, excluded by (R1)).*

Proof. (a) Standard asymptotic MLE theory gives $\sqrt{N_p}(\hat{\theta} - \theta^*) \xrightarrow{d} \mathcal{N}(0, I(\theta^*)^{-1})$, so $\mathrm{Var}_\infty(\hat{J}_{ij}) = [I^{-1}]_{33}/N_p$. Apply Lemma 7.4 with $d = I_{33}$.

The entries of A and b follow from $\mathrm{Cov}(T)$ computed via $\omega_k^2 = 1$: $A_{11} = 1 - m_i^{*2}$, $A_{22} = 1 - m_j^{*2}$, $A_{12} = C_{ij}^* - m_i^* m_j^*$. Positive-definiteness of A on the interior follows from $|C_{ij}^* - m_i^* m_j^*| < \sqrt{(1 - m_i^{*2})(1 - m_j^{*2})}$ away from the boundary of valid Ising distributions.

(b) If $m_i^* = 0$, $b_1 = m_j^*$ and $b_2 = -m_j^* C_{ij}^*$; generically $b \neq 0$ unless $m_j^* = 0$. For $m_i^*, m_j^* \neq 0$: $b = 0$ requires $m_j^* = m_i^* C_{ij}^*$ and $m_i^* = m_j^* C_{ij}^*$, which substituting yields $(C_{ij}^*)^2 = 1$, the degenerate locus.

(c) Solving the linear system explicitly. □

Remark 7.6 (Schur correction: numerical illustration)**.** For $m_i^* = m_j^* = 0.3$, $C_{ij}^* = 0.2$: $b = (0.3 - 0.06, 0.3 - 0.06)^\top = (0.24, 0.24)^\top$, $A = \begin{pmatrix} 0.91 & 0.11 \\ 0.11 & 0.91 \end{pmatrix}$, $A^{-1} \approx \begin{pmatrix} 1.115 & -0.135 \\ -0.135 & 1.115 \end{pmatrix}$, giving $b^\top A^{-1} b \approx 0.0576 \cdot 1.96 \approx 0.113$. Marginal variance: $1/(1 - 0.04) \approx 1.042/N_p$; Schur-corrected efficient variance: $1/(0.96 - 0.113) = 1/0.847 \approx 1.181/N_p$. The correction raises the efficient variance by 13% in this regime, confirming that the marginal bound overstates efficiency when $m_i^* m_j^* \neq 0$.

8 Discussion and Future Work

8.1 Relation to prior Ising-based models of financial markets

The econophysics literature [8, 9, 10, 11, 12, 13, 14] applies Ising-type models to equity markets with J_{ij} modelling trader-imitation strength and ω_i modelling a trader's decision. Derman and Kani [15] model correlation-smile structure in multi-asset derivatives. The present model is different: spins are event outcomes (binary resolutions of real-world events), and J_{ij} is a pairwise interaction term in the *joint distribution over outcomes*, identified from market data on joint contracts. The trader-interaction reading of J is not claimed.

The Ising structure-learning literature provides complementary results. Ravikumar, Wainwright, and Lafferty [18] prove sample-complexity bounds for recovering the sparsity pattern of J from samples of the full spin vector via ℓ_1-regularised logistic regression. Santhanam and Wainwright [29] give the matching information-theoretic lower bound $n = \Omega(c^d \log p)$ for structure recovery, and Bresler [30] gives the first poly(p)-time algorithm without the incoherence condition. Vuffray et al. [19] give a sample-optimal interaction-screening estimator. All of these operate on samples of the full ω-vector and target the full adjacency under sparsity regularisation. The parlay-MLE of Section 3.3 operates on samples of the *pair-marginal* of ω and targets a single edge with unpenalised estimation; the sample-complexity regime is different (parametric $\epsilon^{-2} \log(1/\delta)$ per pair rather than combinatorial scaling in the full-adjacency support). Combining the two, by using market-derived pair-level moment estimates as side information for the ℓ_1 or interaction-screening estimator, is a natural direction.

8.2 Relation to combinatorial prediction markets

Chen and Pennock [16] developed a utility framework for bounded-loss market makers on combinatorial outcome spaces, and Abernethy, Chen, and Wortman Vaughan [17] showed that automated market-making on combinatorial markets admits a convex-optimization formulation. A central design question in that literature is the choice of prior over joint outcomes; the standard LMSR cost function implicitly uses a uniform prior. The Ising parameterisation (3) provides an alternative, namely a prior with pairwise moment constraints, and the MLE of Theorem 3.6 gives a procedure for estimating its parameters from market data. Whether an LMSR with an Ising-parameter prior strictly dominates the uniform-prior LMSR on specific classes of correlated-event markets is an open question.

8.3 Future work: lead-lag-based estimation of J_{ij}

The parlay-based estimator of this paper requires liquid parlay markets for each pair (i, j). In operational practice many correlated-event pairs will have no parlay market, and the analyst may wish to estimate J_{ij} from the lead-lag structure of ordinary marginal-market returns alone. A leading-order bridge between Granger causality statistics on log-return time series and the Ising interaction J_{ij} is available: the population partial-R^2 of the Granger regression is $4J_{ij}^2 + O(J_{\max}^3)$ under the LMSR-to-Boltzmann identification of Remark 2.2, yielding an estimate $|\hat{J}_{ij}| \sim \frac{1}{4}\sqrt{LF_{i\to j}/(W-2L-1)}$ (with L the lag order, W the window, $F_{i\to j}$ the Granger F-statistic). The second-order correction is sensitive to microstructure noise, inventory effects, and third-party couplings, and its sign and magnitude are not determined by the Granger statistic alone. The Granger-based estimator is therefore not suited to the theorem-weight claims of the present paper. A rigorous bridge between marginal-market time-series statistics and the Ising parameters is an open problem; the ℓ_1-regularised approach of [18] applied to binarised returns, or an alternative based on directly observing joint outcomes at event resolutions, are two directions.

8.4 Limitations

The results are stated under explicit modelling assumptions:

(L1) **Pairwise interactions only.** The Ising Hamiltonian (2) cannot represent triple-conjunction correlations or asymmetric tail dependencies. Event triples with joint dependence but pairwise independence (e.g., bracket-conditional tournaments) are misspecified.

(L2) **Stationarity of J^*.** Theorems 3.6 and 4.1 assume the truth is constant over the observation window. Drift in J^* (rule changes, roster shifts, regime changes) requires the confidence intervals on $\hat{J}$ to be inflated by the estimated drift rate.

(L3) **Symmetric $J_{ij} = J_{ji}$.** The Ising parameterisation is undirected. Asymmetric causal relationships cannot be represented; a directed graphical-model alternative is more expressive and lacks the simple closed-form identifiability of Theorem 3.2.

(L4) **LMSR-to-Ising bridge.** The identification $h_i = \frac{1}{2}\operatorname{logit}(p_i)$ assumes quoted marginals equal the Boltzmann marginals. Market-scoring-rule markets satisfy this at equilibrium up to $O(1/b)$ inventory corrections. The robustness theorem of Section 6 bounds the propagation of those perturbations; a full model of the price-formation process mapping the Boltzmann distribution to observed order-book prices is not developed here.

(L5) **Admissible-region boundary.** When a parlay price lies near the Fréchet-Hoeffding boundary, the closed-form estimator is unstable and the MLE has large variance (Remark 6.2). This regime requires a separate treatment.

(L6) **Mean-field convergence condition.** Theorem 2.5 requires $d_{\max} J_{\max} < 1$. In dense networks with strong coupling this condition can fail; the hierarchical / cluster-based extensions noted in the classical literature [13, 14] are beyond the present scope.

(L7) **Survivorship bias.** The estimator operates on the subset of events with sustained trading activity; events whose markets are thinned or de-listed are not observed, and $\hat{J}$ reflects the surviving subset rather than a random sample.

(L8) **Causal sufficiency.** Assumption 7.1(R4) imposes causal sufficiency. Two events that co-move because of an unobserved third factor produce nonzero $\hat{J}$ even when the direct interaction is zero; distinguishing causal from confounded correlations requires additional structure (instrumental variables, natural experiments from resolutions, or the latent-variable identifiability theory of Richardson-Spirtes [36]). The paper's theorems are statements within the pairwise-Ising model under (R4); they do not claim correctness of the model in the presence of latents.

8.5 Open problems

Open Problem 1 (Sample complexity for simultaneous pair selection)**.** Theorem 4.1 gives a per-pair concentration bound. Simultaneously estimating all $\binom{M}{2}$ interactions from a market with N_p parlays on each pair requires a union bound scaling the sample complexity by $O(\log M)$. When only a sparse subset of pairs is relevant (mean network degree D), can adaptive allocation of parlay volume across pairs achieve the oracle rate $N_p \geq C(\theta^*)\epsilon^{-2}\log(DM/\delta)$? The ℓ_1-regularised estimators of [18, 19] answer this for spin-sample observations; the parlay-price version is open.

Open Problem 2 (Identifiability from partial parlay coverage)**.** In practice only a subset $\mathcal{P} \subset \binom{[M]}{2}$ of pairs has liquid parlay markets. Under what graph-theoretic conditions on $\mathcal{P}$ (connectivity, chordality) is the full J matrix identifiable from marginals on all sites plus parlay prices on $\mathcal{P}$? A natural sufficient condition is that the graph $([M], \mathcal{P})$ be connected and contain no cycles of length ≥ 4; tight necessary and sufficient conditions are open.

Open Problem 3 (Optimal parlay-market design for estimation)**.** Given a budget of N_p parlay trades to distribute across $\binom{M}{2}$ pairs, which allocation minimises the weighted sum of estimator variances $\sum_{ij} w_{ij}\mathrm{Var}(\hat{J}_{ij})$? The Fisher information (12) varies across pairs as a function of the local moments (m_i, m_j, C_{ij}); an adaptive allocation that prioritises pairs with small $I_{33} - b^\top A^{-1} b$ (high-variance pairs) would reduce the max-variance bound but requires a prior on θ^*. Formal minimax allocation schemes are an open direction.

Open Problem 4 (Beyond the pairwise model)**.** Higher-order Boltzmann machines with k-spin interactions $J_{i_1 \dots i_k}$ capture genuine joint dependencies that the pairwise Ising model misses. Identification of J_{ijk} requires triple-event parlays (contracts on $\omega_i = \omega_j = \omega_k = +1$); deriving the analogue of Theorem 3.2's admissible region, and the corresponding MLE efficiency, is open. The sample-complexity penalty for k-spin interactions is exponential in k and may be prohibitive beyond $k = 3$.

Open Problem 5 (Non-stationarity)**.** When the true J_t^* drifts at bounded rate $\|J_{t+1}^* - J_t^*\| \leq \delta_J$ per unit time, a stationary MLE of Section 3.3 applied to a window of N_p contiguous parlay observations has bias contribution $O(\delta_J N_p)$ from within-window drift plus MLE variance contribution $O(N_p^{-1/2})$; balancing gives optimal window $N_p^* \sim \delta_J^{-2/3}$ and optimal root-MSE $\sim \delta_J^{1/3}$ (standard rates for tracking a slowly-varying parameter [7]). Deriving the estimator that achieves this trade-off uniformly over specified drift classes (bounded-variation, Lipschitz, or stochastic-process), and the corresponding finite-sample behaviour at pair-specific Fisher information, is open and application-dependent.

References

[1] R. Hanson. Combinatorial information market design. *Information Systems Frontiers*, 5(1):107-119, 2003.

[2] R. Hanson. Logarithmic market scoring rules for modular combinatorial information aggregation. *Journal of Prediction Markets*, 1(1):3-15, 2007.

[3] R. Rana, V. Nadkarni, N. Moshrefi, and P. Viswanath. ParlayMarket: Automated market making for parlay-style joint contracts. arXiv:2603.22596, 2026.

[4] C. Genest and J. V. Zidek. Combining probability distributions: A critique and an annotated bibliography. *Statistical Science*, 1(1):114-135, 1986.

[5] R. A. Fisher. On the "probable error" of a coefficient of correlation deduced from a small sample. *Metron*, 1:3-32, 1921.

[6] E. Ising. Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik*, 31:253-258, 1925.

[7] S. Haykin. *Adaptive Filter Theory*. Pearson, 5th edition, 2014.

[8] J.-P. Bouchaud and R. Cont. Herd behavior and aggregate fluctuations in financial markets. *Macroeconomic Dynamics*, 4(2):170-196, 2000.

[9] J.-P. Bouchaud and R. Cont. A Langevin approach to stock market fluctuations and crashes. *European Physical Journal B*, 6(4):543-550, 1998.

[10] D. Sornette. *Why Stock Markets Crash: Critical Events in Complex Financial Systems*. Princeton University Press, 2003.

[11] T. Kaizoji. Speculative bubbles and crashes in stock markets: an interaction model of speculative activity. *Physica A*, 287(3-4):493-506, 2000.

[12] T. Kaizoji. An interaction model of financial markets from the viewpoint of nonextensive statistical mechanics. *Physica A*, 370(1):109-113, 2006.

[13] M. Mézard, G. Parisi, and M. A. Virasoro. *Spin Glass Theory and Beyond*. World Scientific Lecture Notes in Physics, Vol. 9, 1987.

[14] M. Mézard, G. Parisi, and R. Zecchina. Analytic and algorithmic solution of random satisfiability problems. *Science*, 297:812-815, 2002.

[15] E. Derman and I. Kani. Riding on a smile. *Risk Magazine*, 7(2):32-39, 1994.

[16] Y. Chen and D. M. Pennock. A utility framework for bounded-loss market makers. In *Proceedings of UAI 2007*, pp. 49-56, 2007.

[17] J. Abernethy, Y. Chen, and J. Wortman Vaughan. Efficient market making via convex optimization, and a connection to online learning. *ACM Transactions on Economics and Computation*, 1(2):12:1-12:39, 2013.

[18] P. Ravikumar, M. J. Wainwright, and J. D. Lafferty. High-dimensional Ising model selection using ℓ_1-regularized logistic regression. *Annals of Statistics*, 38(3):1287-1319, 2010.

[19] M. Vuffray, S. Misra, A. Lokhov, and M. Chertkov. Interaction screening: Efficient and sample-optimal learning of Ising models. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2016.

[20] A. Wald. Note on the consistency of the maximum likelihood estimate. *Annals of Mathematical Statistics*, 20(4):595-601, 1949.

[21] W. K. Newey and D. L. McFadden. Large sample estimation and hypothesis testing. In *Handbook of Econometrics*, Vol. 4, Elsevier, pp. 2111-2245, 1994.

[22] R. I. Jennrich. Asymptotic properties of non-linear least squares estimators. *Annals of Mathematical Statistics*, 40(2):633-643, 1969.

[23] E. L. Lehmann and G. Casella. *Theory of Point Estimation*. Springer Texts in Statistics, 2nd edition, 1998.

[24] L. D. Brown. *Fundamentals of Statistical Exponential Families*. IMS Lecture Notes Monograph Series 9, 1986.

[25] B. Efron. *Exponential Families in Theory and Practice*. Cambridge University Press, 2022.

[26] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 2nd edition, 2013.

[27] J. Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B*, 36(2):192-236, 1974.

[28] D. H. Ackley, G. E. Hinton, and T. J. Sejnowski. A learning algorithm for Boltzmann machines. *Cognitive Science*, 9(1):147-169, 1985.

[29] N. P. Santhanam and M. J. Wainwright. Information-theoretic limits of selecting binary graphical models in high dimensions. *IEEE Transactions on Information Theory*, 58(7):4117-4134, 2012.

[30] G. Bresler. Efficiently learning Ising models on arbitrary graphs. In *Proceedings of the 47th ACM Symposium on Theory of Computing (STOC)*, pp. 771-782, 2015.

[31] C. F. Manski. Interpreting the predictions of prediction markets. *Economics Letters*, 91(3):425-429, 2006.

[32] T. Plefka. Convergence condition of the TAP equation for the infinite-ranged Ising spin glass model. *Journal of Physics A: Mathematical and General*, 15(6):1971-1978, 1982.

[33] H. J. Kappen and F. B. Rodríguez. Efficient learning in Boltzmann machines using linear response theory. *Neural Computation*, 10(5):1137-1156, 1998.

[34] Y. M. M. Bishop, S. E. Fienberg, and P. W. Holland. *Discrete Multivariate Analysis: Theory and Practice*. MIT Press, 1975.

[35] A. Agresti. *Categorical Data Analysis*. Wiley, 3rd edition, 2013.

[36] T. Richardson and P. Spirtes. Ancestral graph Markov models. *Annals of Statistics*, 30(4):962-1030, 2002.