

Computation of Collision Local Times: A Connection Between New Scalable Probit Methodology and Atlas Models

Peter Cotton*

September 2026

1 The tie density of a one-shot race

Consider a one-shot race among n contestants with performances $X_1, \dots, X_n$, smallest wins, locations μ and a smooth joint density; $p_i(\mu)$ is i 's win probability. Cotton (2026) (Proposition 1 there, the photo-finish flux) shows that the derivative of i 's win probability with respect to a rival's location is a boundary flux,

$$w_{ij} = \frac{\partial p_i}{\partial \mu_j} = \int \varphi_{ij}(t, t) \Pr\{X_k > t \ \forall k \neq i, j \mid X_i = X_j = t\} dt, \quad (1)$$

the probability density that i and j dead-heat *ahead of the field*, with φ_{ij} the joint density of (X_i, X_j) . We call w_{ij} the *tie density*. The tie densities are symmetric and nonnegative, and because a common shift of every location moves no probability, the Jacobian J they fill has zero row sums: it is the graph Laplacian of the ties. The same reference evaluates p , the tie densities and J 's action for all n contestants together on a lattice of L points in $O(nL)$, by sharing one survival field $\sum_j \log S_j$ across contestants, with S_j the survival function of X_j ; that cost is what makes everything below computable at scale.

2 Brownian races and collision local times

Let the performances be Brownian,

$$X_i(t) = \mu_i + \sigma W_i(t), \quad W_1, \dots, W_n \text{ independent,}$$

so that the time- T marginals form exactly the Gaussian race of Cotton (2026) with idiosyncratic variances $D_i = \sigma^2 T$: the static machinery prices $p_i(T) = \Pr\{X_i(T) = \min_k X_k(T)\}$ and its Jacobian at every horizon, in $O(nL)$ per horizon.

*Research note. Code, experiments and this source live at github.com/microprediction/winning under [research/equilibrium/](https://github.com/microprediction/winning). Companion to Cotton (2026), whose introduction defers this connection “for another day.” Every citation is verified to title, venue, volume and pages against the publisher's locator, but none has yet been read in full; that reading is the next gate before any of this is published.

Now watch the boundary rather than the marginal. Fix a pair (i, j) and set $Y = X_i - X_j$, a Brownian motion with drift $\mu_i - \mu_j$ and quadratic variation $\langle Y \rangle_t = 2\sigma^2 t$. The event “ i and j are the leading pair and tied” is the particle system touching $\{X_i = X_j \leq X_k \ \forall k\}$. Time spent straddling that set is measured by a local time: define the *leading-pair collision local time*

$$\Lambda_T^{ij} = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_0^T \mathbf{1}\{|Y_s| < \varepsilon, X_i(s) \wedge X_j(s) \leq X_k(s) \ \forall k \neq i, j\} d\langle Y \rangle_s.$$

Claim 1. *For the independent Brownian race,*

$$\mathbb{E} \Lambda_T^{ij} = 2\sigma^2 \int_0^T w_{ij}(s) ds = 2\sigma^2 \int_0^T \frac{\partial p_i}{\partial \mu_j}(s) ds, \quad (2)$$

where $w_{ij}(s)$ is the static tie density (1) of the time- s race. Equivalently: the Jacobian entry at horizon s is, up to the factor $2\sigma^2$, the instantaneous rate at which the leading-pair collision local time accrues.

Derivation sketch.

$$\mathbb{E} \Lambda_T^{ij} = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \mathbb{E} \int_0^T \mathbf{1}\{|Y_s| < \varepsilon, (i, j) \text{ lead}\} d\langle Y \rangle_s \quad (3)$$

$$= 2\sigma^2 \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_0^T \Pr\{|Y_s| < \varepsilon, (i, j) \text{ lead}\} ds \quad (4)$$

$$= 2\sigma^2 \int_0^T \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \Pr\{|Y_s| < \varepsilon, (i, j) \text{ lead}\} ds \quad (5)$$

$$= 2\sigma^2 \int_0^T \int \varphi_{ij}^{(s)}(t, t) \Pr\{X_k(s) > t \ \forall k \neq i, j \mid X_i(s) = X_j(s) = t\} dt ds. \quad (6)$$

Step (3) is the definition of Λ_T^{ij} with the expectation brought inside the limit. Step (4) uses $d\langle Y \rangle_s = 2\sigma^2 ds$ and Fubini: the expectation of an indicator is a probability. Step (5) exchanges the limit with the time integral. Step (6) evaluates the band probability to first order in ε : writing $\varphi_{ij}^{(s)}$ for the joint density of $(X_i(s), X_j(s))$,

$$\begin{aligned} \Pr\{|Y_s| < \varepsilon, (i, j) \text{ lead}\} &= \iint_{|u-v| < \varepsilon} \varphi_{ij}^{(s)}(u, v) \Pr\{\text{field above } u \vee v \mid u, v\} du dv \\ &= 2\varepsilon \int \varphi_{ij}^{(s)}(t, t) \Pr\{\text{field above } t \mid t, t\} dt + o(\varepsilon), \end{aligned}$$

the strip of width 2ε around the diagonal collapsing onto it. The inner integral of (6) is precisely (1) for the time- s race, i.e. $w_{ij}(s)$. What needs care to make this a proposition: the ε -limit exchange (dominated convergence with the continuous conditional density), and the leading-pair indicator, whose discontinuity set is the triple-tie event — measure zero for non-degenerate diffusions, and exactly the degeneracy the Atlas literature attends to (Ichiba and Karatzas, 2010). Correlated (factor) versions keep the time- T marginals inside the covariance grammar of Cotton (2026); the identity needs only $\langle Y \rangle$ adjusted for the factor covariance — mechanical, unwritten. $\square$

So the boundary flux is not an *analogy* to local time: for Brownian races it *is* the local-time density, and the shared field computes it — implicitly for all $\binom{n}{2}$ pairs, and explicitly for any one pair — at $O(nL)$ per horizon.

3 Numerical checks

Three checks, all seeded and committed under `research/equilibrium/exp2_localtime` at `github.com/microprediction/winning`.

Claim 1 against simulation. A five-runner Brownian race ($\sigma = 1$, $T = 0.5$, distinct locations), 60,000 exactly simulated paths at $dt = 2.5 \times 10^{-5}$, occupation estimator run at band widths $\varepsilon = 0.08$ and 0.04 and extrapolated linearly to $\varepsilon \rightarrow 0$ (the finite-band bias is visibly linear: the raw estimates sit 2.7–4.4 percent low). Against the lattice side of (2), stable under doubling of the horizon grid: the two favourites, Monte Carlo 0.3539 ± 0.0038 against lattice 0.3544 ($z = -0.14$); the favourite against a midfield runner, 0.0946 ± 0.0022 against 0.0956 ($z = -0.47$).

The trace identity. Summing over named pairs (i, j) gives the local time of the gap between the ranked first and second particles, whoever they are, and J 's zero row sums make the total a trace: $\sum_{i < j} w_{ij} = -\frac{1}{2} \text{tr } J$. Measured identical to six decimals (0.995801 both ways). The trace is the sum of the own-coordinate derivatives $\partial p_i / \partial \mu_i$, which the forward pass of Cotton (2026) returns with the probabilities, so the whole-field collision rate is free with pricing.

Scale. For a field the size of the Russell 3000 index, $n = 3000$, the full expected rank-1/rank-2 collision-rate curve across 24 horizons computes in 1.43 seconds on one laptop. Collision local times at this scale are not otherwise available by any method we know.

4 Calibration to market weights at scale

The identities above become *inference* through the calibration of Cotton (2026). Market weights are nonnegative and sum to one, so a cross-section of weights is a choice-probability vector, and the inversion problem — find locations μ for which the race reproduces an observed vector — is Theorem 1 of that reference: for any interior target the inverse exists and is unique up to a common shift, as the maximizer of a strictly concave program whose Hessian is J itself — minus the graph Laplacian of Section 1, negative definite on the contrast space. The shipped solver is a damped diagonal iteration on the log residual, preconditioned by the same own-coordinate derivatives that the trace identity uses, and the reported round trips run at $n = 10^6$: about a minute for independent performances, about twenty minutes for a rank-one factor, with maximum log-probability residuals below 10^{-8} . Forward pricing alone at $n = 10^4, 10^5, 10^6$ takes 0.18, 2.7 and 29 seconds.

Correlation enters through a covariance grammar priced exactly: a global factor ($VV' + \text{diag } D$), blocks (a rank- r shared effect per cluster — sectors), nested (blocks under a global factor — market plus sectors), and trees of uniform shared effects. The tree case is the one an equity reader should notice: Cotton (2026) exhibits a tree race whose implied correlation reproduces the floored cophenetic correlation of a linkage dendrogram exactly, so the hierarchical correlation structure behind hierarchical risk parity sits inside the grammar, and a stock field can be calibrated under sector, factor or dendrogram covariance rather than independence. Dense covariance outside the grammar is fitted approximately, with the fit residual reported per named random-correlation ensemble.

The consequence for this note: every local-time quantity above can be evaluated at the calibrated point. Fit a covariance (sectors, a dendrogram, a factor model), invert today's

weights to drifts, and the collision-rate curves and the leakage rates of the next section become *market-implied* — implied in the sense of implied volatility, the values that make the model reproduce the observed cross-section — with the whole pipeline running in minutes at the scale of the Russell 3000.

5 Index leakage

Fernholz calls the relative-return loss a capitalization-league portfolio suffers when stocks fall out of the league and are replaced *leakage*, and expresses it through semimartingale local time at the dividing rank (Fernholz, 2002). Index entry and exit — reconstitution turnover — is a rank-boundary crossing, so the machinery above prices it, once the boundary moves from rank one to rank k . The tie density generalizes cleanly: for $q_i^{(k)} = \Pr\{X_i \text{ among the } k \text{ smallest}\}$ the same divergence-theorem argument puts the whole derivative on ties of i and j *at the cutoff*,

$$\frac{\partial q_i^{(k)}}{\partial \mu_j} = \int \varphi_{ij}(t, t) \Pr\{\#\{m \neq i, j : X_m < t\} = k - 1 \mid X_i = X_j = t\} dt, \quad (7)$$

with $k = 1$ recovering (1). The count probability is Poisson–binomial over the survivals the field already evaluates, $O(nLk)$ by the standard recursion. Summed over pairs and scaled by $2\sigma^2$, (7) is the expected local-time rate at the rank- $(k, k + 1)$ gap: the leakage rate of a top- k index under a fitted race model, at Russell scale ($k = 1000$ or 3000) at the same cost orders as everything above. Implementation is tracked as issue 12 of the `winning` repository; equation (7) is derived here and not yet implemented or measured.

6 Why the Atlas literature should care

Rank-based (Atlas) models (Fernholz, 2002; Banner et al., 2005) assign drift and volatility by *rank*; the semimartingale decomposition of the k -th ranked process carries the collision local times $\Lambda^{(k, k+1)}$ between adjacent ranks as correction terms, and in stochastic portfolio theory the size effect and the attrition of the largest market weight are expressed through them. Market weights are a choice-probability vector — the portfolio remark of Cotton (2026) is the same bridge — and rank-crossing local times are what the static machinery evaluates, one horizon at a time.

Trade flows in both directions.

1. *Computation to Atlas.* Collision local times are analytically available only for small n or special symmetric cases. Claim 1 plus the lattice evaluates expected local-time densities for the leading pair at $O(nL)$ per horizon for n in the millions; the removal counterfactual of Cotton (2026) (delete the leaders, re-race) reaches lower ranks. Checkable against simulated ranked Brownian systems — an exact-versus-Monte-Carlo experiment of the kind that reference’s benchmarks already referee.
2. *Structure to us.* The Atlas decomposition names the dynamic inversion: recover drifts from observed *leadership turnover* (occupation and crossing rates of ranks), of which the static inversion is the one-horizon shadow.

7 Limitations

The verified identity is for the leading pair; the rank- k boundary formula (7) is derived but unimplemented and unmeasured. Market-implied local times inherit the covariance grammar: a mis-specified covariance moves the implied rates, and nothing here measures that sensitivity. Nothing here addresses stationary gap distributions (Pal and Pitman, 2008), where much of the Atlas literature lives; the contribution in prospect is finite-horizon and computational. Citations are locator-verified only: the sources have not yet been read in full, and in particular whether the leading-pair (rather than adjacent-rank) local time already carries a name in that literature is unchecked.

References

- A. D. Banner, R. Fernholz, I. Karatzas. Atlas models of equity markets. *Annals of Applied Probability* 15(4):2296–2330, 2005.
- P. Cotton. Scalable inversion of contests with correlated performances, including softmax and multinomial probit. Working paper, arXiv (submitted), 2026. Code and live demonstration: winning.microprediction.org.
- E. R. Fernholz. *Stochastic Portfolio Theory*. Springer, New York, 2002.
- T. Ichiba, I. Karatzas. On collisions of Brownian particles. *Annals of Applied Probability* 20(3):951–977, 2010.
- S. Pal, J. Pitman. One-dimensional Brownian particle systems with rank-dependent drifts. *Annals of Applied Probability* 18(6):2179–2207, 2008.