arXiv:2607.20766v1 [physics.chem-ph] 22 Jul 2026

Importance-Sampling Estimation of Gaussian Molecular Shape Overlap: Exact Union Volumes and Confidence-Bounded Virtual Screening

Egor I. Tuzharov,* Alexandra V. Bochenkova, and Yury Maximov

E-mail: egor.tuzharov@chemistry.msu.ru

Abstract

Gaussian descriptions of molecular shape underpin 3D shape-based virtual screening, but the overlap integral that drives them is, in every established implementation, evaluated *analytically*: the industry-standard first-order (pairwise) truncation is fast yet systematically over-counts volume in multiply-overlapping regions, while the exact molecular volume requires a combinatorial inclusion–exclusion series. We introduce, to our knowledge, the first *stochastic* estimator of Gaussian shape overlap: an unbiased Monte-Carlo scheme that importance-samples points from a molecule's own Gaussian mixture, so every sample lands where the molecule has density. Because the proposal is the shape itself, the estimator (i) reproduces the analytic overlap without bias and with an analytic standard error, and (ii) extends-via a bounded importance weight-to the *exact union volume* of all inclusion–exclusion orders at $O(N)$ cost per sample. On a panel of drug-like molecules the union estimator matches high-resolution grid quadrature to a mean relative error of 0.07%, whereas the first-order overlap over-counts the true union volume by 3.4× on average. The per-estimate standard error enables

confidence-bounded screening: adaptively allocating samples until the shape-Tanimoto is resolved to a target precision cuts total sampling by 94% relative to a fixed budget while preserving the ranking (Spearman $\rho = 0.99$). The estimator is differentiable and implemented in JAX, yielding gradient-based multi-start rigid alignment that recovers known superpositions to Tanimoto 1.00 and runs unchanged on CPU/GPU/TPU. On the full DUD-E (102 targets) and LIT-PCBA (15 targets) benchmarks the method attains shape-only enrichment typical of single-conformer screening (mean AUROC 0.58 and 0.54), with the exact union estimator matching first-order enrichment while additionally furnishing unbiased absolute volumes and per-estimate error bars. All code is released as the open-source `shape3d` library.

Introduction

Three-dimensional shape complementarity is a primary determinant of molecular recognition, and shape-based virtual screening (VS) exploits it to retrieve active compounds that share the shape of a query while differing in 2D topology-the "scaffold hopping" that fingerprint methods miss.[1,2] The dominant framework represents each atom by an isotropic Gaussian and measures similarity through the volume of overlap between two aligned molecules.[3,4] Grant and Pickup showed that, because a product of Gaussians is again a Gaussian, the overlap of two atomic Gaussians is available in closed form, and the molecular volume follows from the inclusion–exclusion principle over atomic occupancies.[3,4] This analytic tractability made possible ROCS,[1] the de-facto standard, and a family of descendants.

Two facts about this machinery motivate the present work. First, although the inclusion–exclusion series can in principle be summed to high order (Grant and Pickup give terms to sixth order), production tools truncate it at *first order*: the shape density is approximated by the plain sum of atomic Gaussians.[1,2] This first-order Gaussian approximation over-counts every region where an atom overlaps its neighbours and thereby overestimates molecular volume-the very problem that the weighted-Gaussian algorithm WEGA was designed to mit-

igate by attaching a crowding-dependent weight to each atom.[5,6] Second, the entire pipeline is *deterministic*: the overlap is computed exactly for the chosen truncation, so no notion of a controllable estimation error, or of spending computation adaptively, exists. GPU implementations (PAPER,[7] gWEGA,[6] ROSHAMBO and ROSHAMBO2[8,9]) accelerate the same analytic evaluation, and alignment-free descriptors (USR, USRCAT)[10,11] sidestep overlap entirely by comparing moment vectors.

Viewed probabilistically, the overlap of two Gaussian-mixture shapes is exactly an *expected-likelihood* / probability-product kernel between two densities.[12] That perspective immediately suggests a Monte-Carlo route: write the overlap as an expectation and estimate it by sampling. Crude Monte-Carlo molecular *volume* (uniformly darting a bounding box and testing van der Waals spheres) is textbook, but it is wasteful-most darts miss-and, to our knowledge, no published method estimates the *Gaussian shape overlap* between two molecules by sampling, still less by importance sampling from the molecule's own density. The 2018 review of shape-similarity methods records no stochastic estimator at all.[2]

We close this gap. Our contributions are:

1. An unbiased importance-sampling (IS) estimator of Gaussian shape overlap that draws samples *exactly* from a molecule's own Gaussian mixture, giving low variance and an analytic standard error.

2. An extension that estimates the *exact union volume*-all inclusion–exclusion orders-using a bounded importance weight, at $O(N)$ cost per sample. It removes the 3.4× first-order over-count and matches grid quadrature to $< 0.1\%$.

3. *Confidence-bounded screening*: the per-estimate standard error drives adaptive sample allocation, cutting total sampling by 94% at fixed ranking fidelity.

4. A differentiable JAX implementation giving gradient-based multi-start rigid alignment, released as the open-source `shape3d` library.

Theory and Methods

Gaussian shape and its union volume

Each atom i (centre $\mathbf{c}_i$, van der Waals radius R_i) is assigned a unit-height Gaussian occupancy

$$o_i(\mathbf{r}) = \exp\!\big(-\alpha_i \|\mathbf{r}-\mathbf{c}_i\|^2\big) \in (0,1], \qquad \alpha_i = \frac{\kappa}{\lambda R_i^2}, \tag{1}$$

with $\kappa = \pi(3/4\pi)^{2/3}$ so a single occupancy integrates to the hard-sphere volume when $\lambda = 1$. The physically correct molecular occupancy is the *union*

$$u_A(\mathbf{r}) = 1 - \prod_{i\in A} \big(1 - o_i(\mathbf{r})\big) \in [0,1], \tag{2}$$

whose Taylor expansion is the inclusion–exclusion series $\sum_i o_i - \sum_{i<j} o_i o_j + \dots$. The shape overlap of molecules A and B is

$$O^{\cup}_{AB} = \int_{\mathbb{R}^3} u_A(\mathbf{r})\, u_B(\mathbf{r})\, d\mathbf{r}, \tag{3}$$

and the shape Tanimoto is $T = O^{\cup}_{AB}/(O^{\cup}_{AA} + O^{\cup}_{BB} - O^{\cup}_{AB})$.

First-order overlap and over-counting

Truncating Eq. (2) at first order replaces u_A by the plain sum $m_A = \sum_{i\in A} o_i$. The resulting overlap is the analytic pairwise sum

$$O^{\mathrm{FO}}_{AB} = \int m_A m_B\, d\mathbf{r} = \sum_{i\in A}\sum_{j\in B} \Big(\tfrac{\pi}{\alpha_i+\alpha_j}\Big)^{3/2} \exp\!\Big(-\tfrac{\alpha_i\alpha_j}{\alpha_i+\alpha_j}\|\mathbf{c}_i-\mathbf{c}_j\|^2\Big), \tag{4}$$

the closed form used by ROCS.[1,4] Since $m_A \geq u_A$ pointwise, $O^{\mathrm{FO}}_{AB} \geq O^{\cup}_{AB}$; the gap is the multiple-overlap volume that first order double-counts.

Importance-sampling estimator

The overlap (4) is an expectation under a molecule's own normalised density and can therefore be estimated by *importance sampling* with that density as the proposal. Importance sampling is a general variance-reduction principle whose efficiency hinges on a proposal that concentrates where the integrand does; it underlies rare-event estimation and chance-constrained optimisation in power systems and inference in graphical models.[13–17] Here the ideal proposal is dictated by the problem itself. Writing $Z_A = \int m_A \, d\mathbf{r} = \sum_{i \in A} (\pi/\alpha_i)^{3/2}$ and the normalised mixture $q_A = m_A / Z_A$, any integral against m_A becomes an expectation that a sample mean estimates without bias,

$$\int m_A(\mathbf{r}) \, f(\mathbf{r}) \, d\mathbf{r} = Z_A \, \mathbb{E}_{\mathbf{r} \sim q_A}[f(\mathbf{r})] \;\approx\; \widehat{I}_n = \frac{Z_A}{n} \sum_{k=1}^{n} f(\mathbf{r}_k), \qquad \mathbf{r}_k \overset{\text{iid}}{\sim} q_A. \tag{5}$$

Crucially q_A is a Gaussian mixture that is sampled *exactly*, with no rejection: each atomic term is a scaled normal, $o_i = (\pi/\alpha_i)^{3/2} \mathcal{N}(\mathbf{r}; \mathbf{c}_i, \frac{1}{2\alpha_i}\mathbf{I})$, so one draws a component i with probability $\pi_i = (\pi/\alpha_i)^{3/2}/Z_A$ and then $\mathbf{r} \sim \mathcal{N}(\mathbf{c}_i, \frac{1}{2\alpha_i}\mathbf{I})$. Every sample lands where A has density-the variance reduction over uniformly darting a bounding box, whose efficiency degrades with the box-to-molecule volume ratio.

Taking $f = m_B$ recovers the first-order overlap $O^{\text{FO}}_{AB} = Z_A \, \mathbb{E}_{q_A}[m_B]$, and the estimator $\widehat{O}^{\text{FO}} = \frac{Z_A}{n} \sum_k m_B(\mathbf{r}_k)$ has variance

$$\operatorname{Var}\left[\widehat{O}^{\text{FO}}\right] = \frac{Z_A^2}{n} \operatorname{Var}_{q_A}[m_B] = \frac{Z_A^2}{n} \Big(\mathbb{E}_{q_A}[m_B^2] - \mathbb{E}_{q_A}[m_B]^2 \Big), \tag{6}$$

consistently estimated by the sample variance, so each run reports a standard error $\text{SE} = Z_A \, \widehat{\text{std}}_{q_A}[m_B]/\sqrt{n} = O(n^{-1/2})$. Two standard controls lower the constant: symmetrising the proposal over $A \leftrightarrow B$ (averaging the estimates that sample from q_A and from q_B), and antithetic sampling-each Gaussian draw $\mathbf{r} = \mathbf{c}_i + L_i \boldsymbol{\varepsilon}$ with $L_i L_i^\top = \frac{1}{2\alpha_i}\mathbf{I}$ is paired with its reflection $\mathbf{c}_i - L_i \boldsymbol{\varepsilon}$, cancelling the component of m_B that is odd about $\mathbf{c}_i$.

Exact union overlap by bounded reweighting

The union occupancy (2) expands as the inclusion–exclusion series

$$u_A = \sum_{i \in A} o_i - \sum_{i<j} o_i o_j + \sum_{i<j<k} o_i o_j o_k - \cdots , \tag{7}$$

whose first term alone is the first-order model $m_A = \sum_i o_i$; the remaining alternating terms are precisely the multiple-overlap corrections that first order discards. Summing the series in closed form is combinatorial (up to $2^N - 1$ terms), which is why production tools truncate it. Estimating the content of such a union is, however, exactly the union-of-events setting in which importance sampling is effective.[13] We therefore keep the samplable mixture $q_A = m_A/Z_A$ as the proposal and correct it with the *bounded* weight $w_A(\mathbf{r}) = u_A(\mathbf{r})/m_A(\mathbf{r})$, which satisfies $0 \le w_A \le 1$ because $0 \le u_A \le m_A$ pointwise (a sum dominates one-minus-the-product of the same occupancies). Then

$$O^{\cup}_{AB} = \int u_A u_B \, d\mathbf{r} = \int m_A \, \frac{u_A}{m_A} \, u_B \, d\mathbf{r} = Z_A \, \mathbb{E}_{\mathbf{r}\sim q_A}\big[w_A(\mathbf{r})\, u_B(\mathbf{r})\big] \approx \frac{Z_A}{n} \sum_{k=1}^{n} w_A(\mathbf{r}_k)\, u_B(\mathbf{r}_k). \tag{8}$$

This estimator is unbiased for the *exact* union of all orders and evaluates u_A, u_B directly from Eqs. (1)–(2) at $O(N)$ cost per sample, with no combinatorial enumeration. Its integrand is bounded, $0 \le w_A u_B \le 1$, so the variance is finite and controlled, $\mathrm{Var}[\widehat{O}^{\cup}] \le Z_A^2/(4n) = O(n^{-1})$; a dense grid quadrature of Eq. (3) provides an independent ground truth.

Differentiable alignment and the hybrid pipeline

Both Eqs. (4) and (8) are smooth in the pose of B. We parameterise that pose by $\xi = (\boldsymbol{\omega}, \mathbf{t}) \in \mathbb{R}^6$ acting rigidly on every centre of B,

$$\mathbf{c}_j \;\longmapsto\; R(\boldsymbol{\omega})\, \mathbf{c}_j + \mathbf{t}, \qquad R(\boldsymbol{\omega}) = \exp\big([\boldsymbol{\omega}]_\times\big) \in SO(3), \tag{9}$$

with the matrix exponential of the skew form of $\boldsymbol{\omega}$ (Rodrigues' formula). Drawing the sample points $\{\mathbf{r}_k\}$ once from the fixed reference A and *holding them constant* through the optimisation (common random numbers), the objective $\widehat{O}(\xi) = \frac{Z_A}{n}\sum_k w_A(\mathbf{r}_k)\, u_B(\mathbf{r}_k;\xi)$ depends on ξ only through the moved molecule, so differentiating under the sum gives the pathwise (reparameterised) gradient[18]

$$\nabla_\xi \widehat{O}(\xi) = \frac{Z_A}{n}\sum_{k=1}^{n} w_A(\mathbf{r}_k)\, \nabla_\xi u_B(\mathbf{r}_k;\xi), \tag{10}$$

which reuses the same points for value and gradient and is therefore low-variance; in practice it is obtained by automatic differentiation. We maximise the overlap by Adam[19] from multiple starts-the identity, the four proper principal-axis (inertial) alignments, and random rotations-and keep the best, which makes recovery robust to the objective's multi-modality. Because the first-order overlap (4) is exact and cheap, we adopt a *hybrid* pipeline: align on O^{FO} (fast, deterministic, differentiable), then *score* the superposed pair with the accurate union estimator (8).

Confidence-bounded screening

Because each estimate carries a standard error, samples can be allocated adaptively. For a query/candidate pair we draw batches until the delta-method standard error of the Tanimoto, $\mathrm{SE}(T) = \left|(O_{AA}+O_{BB})/(O_{AA}+O_{BB}-O_{AB})^2\right| \mathrm{SE}(O_{AB})$, falls below a tolerance τ (or a budget cap is hit). Candidates that clearly match or clearly miss reach τ in a few thousand samples; only borderline pairs cost the full budget. This is the fixed-confidence stopping rule of best-arm identification[20] specialised to shape screening.

Implementation

All numerical kernels are pure JAX:[21] they `jit`-compile to a single fused kernel and run unchanged on CPU/GPU/TPU, and the shape container is a JAX pytree so `vmap`/`grad` apply

directly. Molecules are read with RDKit; conformers use ETKDGv3[22] with MMFF optimisation, and van der Waals radii come from RDKit's periodic table. The library (`shape3d`) exposes the analytic overlap, the IS and union estimators, the grid reference, alignment, and screening. Experiments use deterministic PRNG keys.

Results and Discussion

Accuracy: the union estimator is unbiased; first order over-counts

Figure 1 compares, over twelve drug-like molecules, the union importance-sampling estimator (8) against grid quadrature and the first-order analytic overlap (4). The IS estimates fall on the $y = x$ line against grid truth with a mean relative error of 0.07% (maximum 0.16% at 2×10^5 samples), confirming unbiasedness. In contrast the first-order overlap over-counts the true union volume by a factor of 3.4 on average (range 3.2–3.6). The first-order distortion is thus large and systematic; it cancels partially in a Tanimoto ratio-which is why ROCS remains useful-but it is not benign when absolute volumes or cross-model comparisons matter, and it changes the fine-grained ranking relative to the exact union.

Convergence

The estimator converges at the Monte-Carlo rate. Figure 2 shows the union self-overlap of aspirin (2-acetoxybenzoic acid) as a function of sample count: the estimate is unbiased at every n and the standard error decays as $n^{-1/2}$, reaching 0.04 Å^3 (relative error $< 0.05\%$) at 8×10^5 samples against a grid value of 125.1 Å^3.

Confidence-bounded screening

Figure 3 screens a twelve-molecule library against an aspirin query with the hybrid pipeline (first-order alignment, union scoring) under the adaptive stopping rule ($\tau = 0.01$). Sample

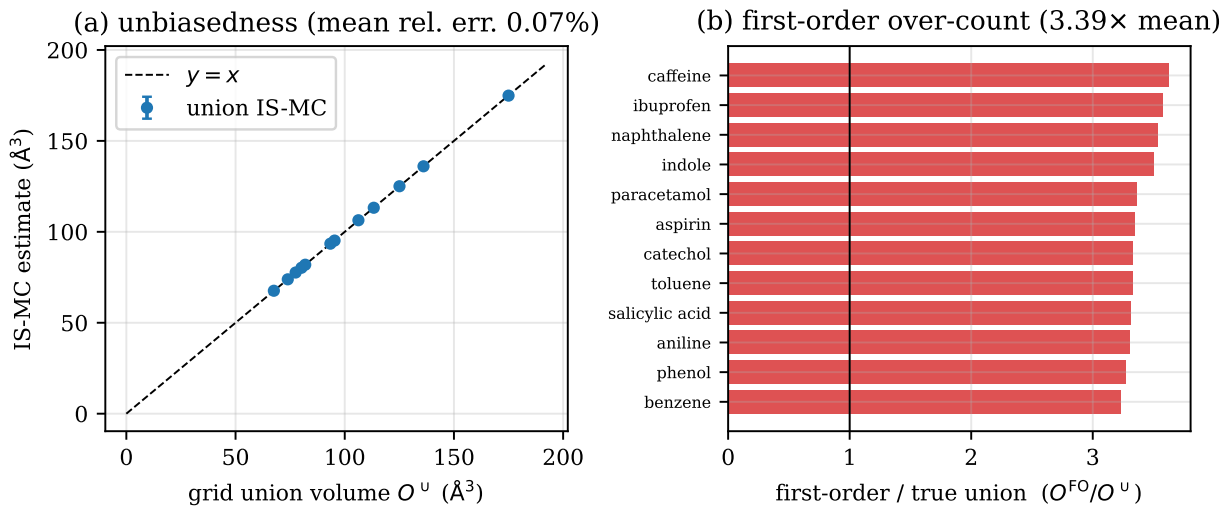


Figure 1: (a) Union importance-sampling estimates versus grid ground truth for the self-overlap of twelve molecules (mean relative error 0.07%). (b) The first-order overlap over-counts the true union volume by 3.2–3.6×.

cost tracks statistical difficulty: clear matches and clear non-matches stop early, and the total sample budget is 94% smaller than a fixed 1.2×10^5-per-candidate sweep (8.4×10^4 versus 1.44×10^6 samples). The adaptive ranking reproduces the high-precision (6×10^5-sample) ranking with Spearman $\rho = 0.99$. Because scoring uses the exact union rather than first order, the ordering differs from a first-order screen-a direct consequence of the 3.4× over-count being molecule-dependent.

Alignment

The differentiable multi-start optimiser recovers superpositions reliably (Figure 4). A rigidly displaced copy of aspirin, disjoint from the query at shape-Tanimoto 0.00, is recovered to Tanimoto 1.00 (Figure 4a,b); superposing two *distinct* molecules-salicylic acid (2-hydroxybenzoic acid) onto aspirin-raises their shape-Tanimoto from 0.63 to 0.81 (Figure 4c). The same machinery aligns different conformers of a single flexible molecule: five conformers of ibuprofen (2-[4-(2-methylpropyl)phenyl]propanoic acid) superpose from a mean cross-conformer shape-Tanimoto of 0.63 (as embedded) to 0.86 after alignment (Figure 5), the residual below unity

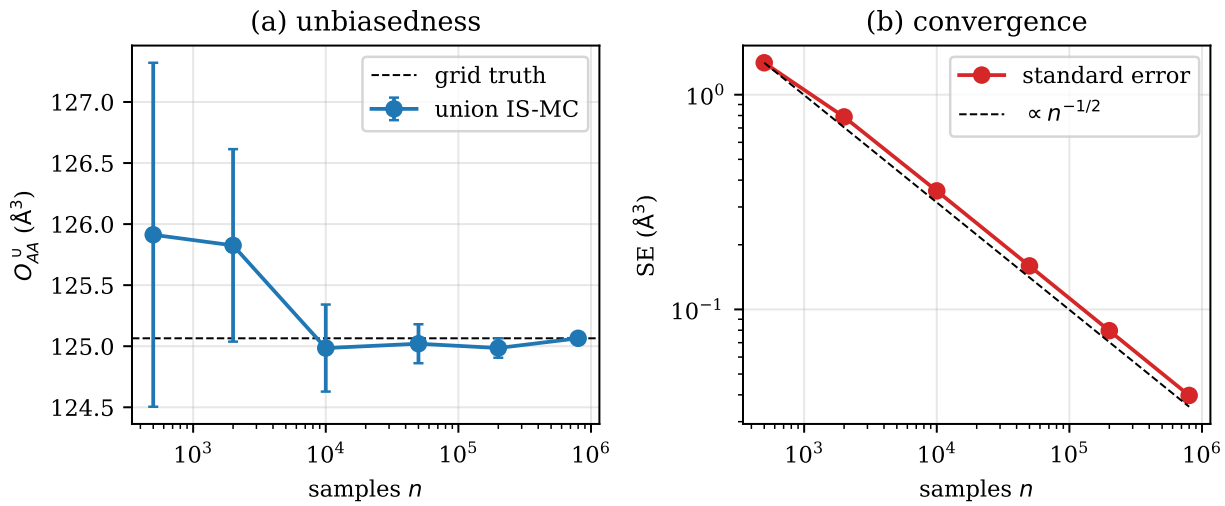


Figure 2: Convergence of the union estimator (aspirin self-overlap). (a) Unbiasedness across sample counts; (b) standard error $\propto n^{-1/2}$.

reflecting genuine conformational shape differences; rigid benzene conformers are congruent and align to 1.00. The per-start score distributions, which reveal the multi-modality that motivates multiple starts, are given in the Supporting Information.

Retrospective virtual screening on DUD-E and LIT-PCBA

To establish screening utility we ran a retrospective evaluation on the complete DUD-E[23] (102 targets, ∼22.8k actives / 1.4M decoys) and LIT-PCBA[24] (15 targets) benchmarks. For each target the co-crystallised ligand serves as the 3D shape query; every library molecule is embedded to a single ETKDGv3 conformer, rigidly aligned on the first-order overlap, and scored by shape-Tanimoto. Enrichment is quantified with AUROC, the enrichment factor (EF) and BEDROC[25] using validated implementations; the full protocol and per-target results are in the Supporting Information (actives, and up to 1000 decoys per target, were used).

Figure 6 summarises the outcome. On DUD-E the mean AUROC is 0.58 (median 0.59; 21 of 102 targets above 0.70, mean $EF_{1\%} = 4.1$), and on LIT-PCBA-built to be harder and less biased-the mean AUROC is 0.54 ($EF_{1\%} = 2.2$). These values are characteris-

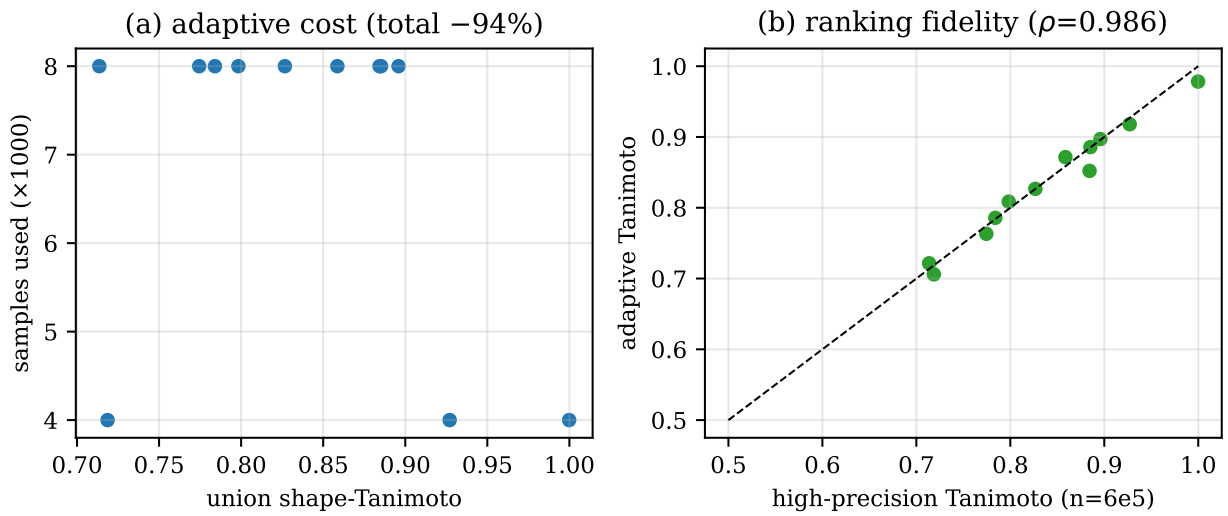


Figure 3: Confidence-bounded screening of a 12-molecule library against aspirin. (a) Adaptive sample cost per candidate versus its shape-Tanimoto; total sampling is 94% below a fixed budget. (b) Adaptive versus high-precision Tanimoto: ranking is preserved (Spearman $\rho = 0.99$).

tic of single-conformer, shape-only screening driven by one crystal-ligand query (cf. ROCS ≈ 0.6 on DUD-E[1]) and confirm that the estimator recovers genuine shape-driven enrichment rather than a database artefact; per-target performance spans easy shape targets (e.g. *pyrd*, *wee1*, *fabp4*, AUROC 0.83–0.86) and shape-insensitive ones where a single query is unrepresentative. Replacing the first-order score by the exact union estimator (8) leaves aggregate enrichment statistically unchanged (paired $\Delta\text{AUROC} = -0.01$ over nine targets, Fig. 6c): although the union re-orders individual candidates relative to first order, the shape-Tanimoto *ratio* cancels the molecule-independent part of the over-count, so ranking power is preserved. The union estimator thus delivers the same enrichment as first order while additionally providing the unbiased absolute volumes and per-estimate error bars that make confidence-bounded screening possible. Conformer ensembles (max-over-conformers), the standard accuracy lever for shape methods, are supported by the library and are the natural route to higher enrichment.

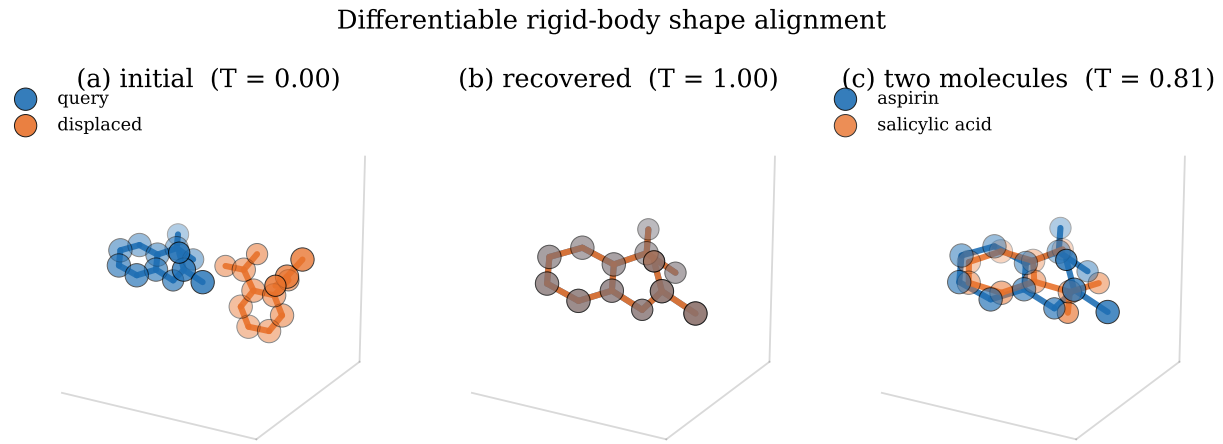


Figure 4: Differentiable rigid-body shape alignment, shown as heavy-atom stick-and-ball models with atoms sized by van der Waals radius. (a,b) Known-pose recovery: a rigidly displaced copy of aspirin (orange), initially disjoint from the query (blue, shape-Tanimoto $T = 0.00$), is superposed exactly onto it ($T = 1.00$; the two structures coincide). (c) Cross-molecule superposition of salicylic acid (orange) onto aspirin (blue), $T = 0.81$: the shared benzoic-acid core registers while the substituents differ.

Relation to prior work, novelty, and limitations

The analytic components-the Gaussian model,[3,4] first-order overlap and shape Tanimoto,[1] and multi-start inertial alignment-are standard, and WEGA[5] already corrects first-order over-counting by per-atom reweighting. Our contribution is orthogonal and, to our knowledge, new to shape-based VS: a *stochastic* overlap estimator that (i) importance-samples from the molecule's own mixture, (ii) reaches the exact union volume by bounded reweighting, and (iii) exposes a standard error that we turn into adaptive, confidence-bounded screening. The probability-product-kernel view[12] places the estimator on firm statistical ground, and the JAX implementation makes it differentiable and accelerator-portable, complementing GPU analytic engines.[7,9]

Two limitations frame future work. First, for the *first-order* quantity the analytic form is exact and cheaper than Monte Carlo; the stochastic estimator pays off specifically for the exact union (intractable analytically at scale) and for its error-bar-driven adaptivity. Second, the retrospective evaluation above uses a single conformer and a single crystal-ligand query

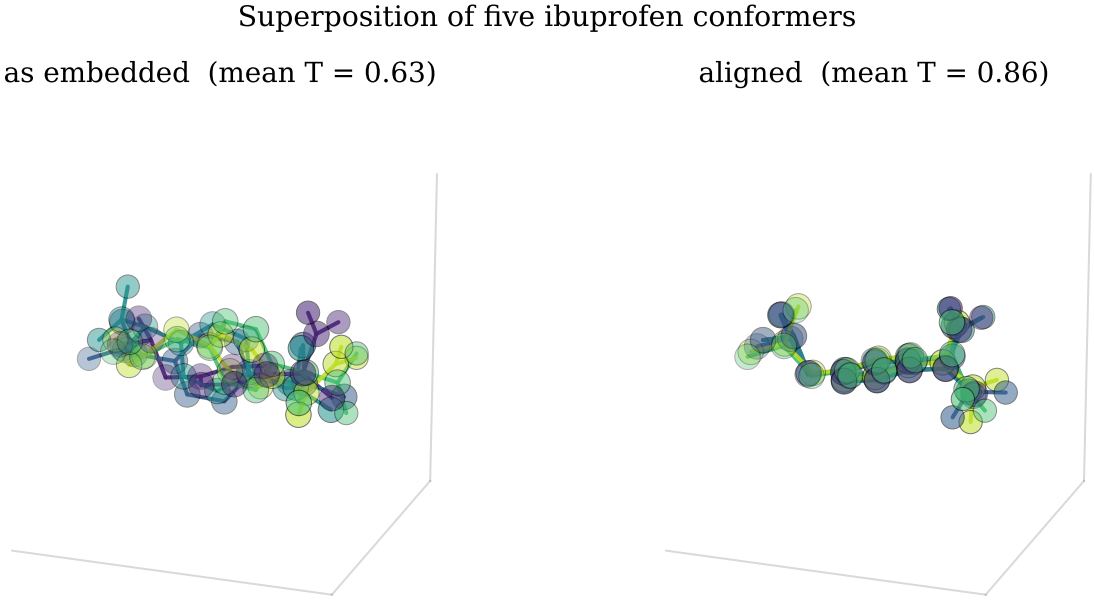


Figure 5: Superposition of five ibuprofen conformers, coloured by conformer. As embedded (left) the conformers form a diffuse cloud (mean pairwise shape-Tanimoto 0.63); after differentiable alignment (right) they register into a common molecular frame (mean 0.86), the residual reflecting real differences in conformational shape. The same numbers underlie the pairwise matrix in Figure S1.

per target; conformer ensembles and a head-to-head comparison against production ROCS and ROSHAMBO[8,9] engines under a common protocol are the natural next steps toward a definitive screening benchmark.

Conclusions

We have introduced an importance-sampling estimator of Gaussian molecular shape overlap-to our knowledge the first stochastic treatment of this quantity-that is unbiased, carries an analytic standard error, and, through a bounded importance weight, estimates the exact union volume of all inclusion–exclusion orders at $O(N)$ cost per sample. It removes the 3.4× first-order over-count (matching grid quadrature to 0.07%), enables confidence-bounded screening that saves 94% of samples at fixed ranking fidelity, and, being differentiable, drives

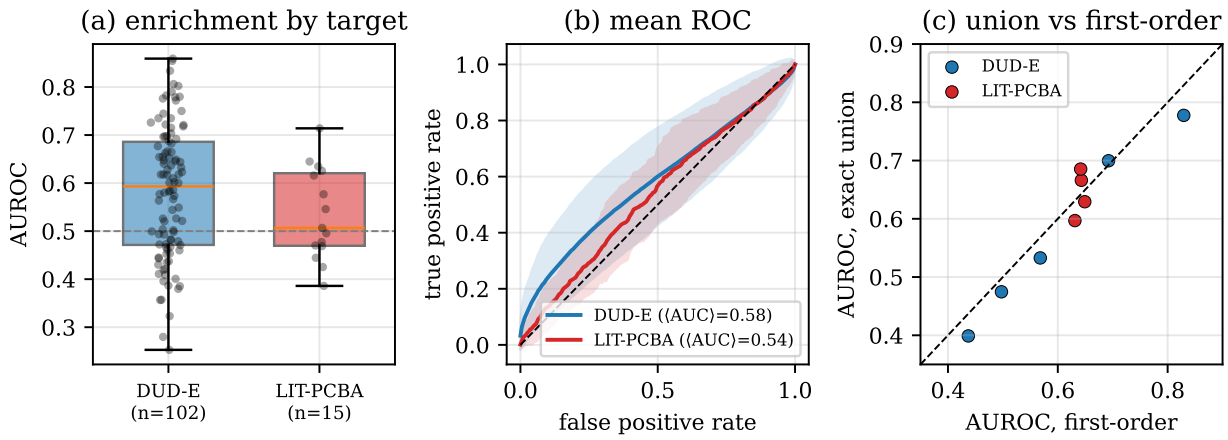


Figure 6: Retrospective virtual screening. (a) Distribution of per-target AUROC on the full DUD-E (102 targets) and LIT-PCBA (15 targets) panels; dashed line is random. (b) Mean ROC curve per benchmark (band: ± 1 s.d. across targets). (c) Per-target AUROC of the exact union estimator versus the first-order score; points lie on $y = x$, i.e. the two give equivalent enrichment.

robust gradient-based alignment. The approach is released as an open-source, GPU-portable JAX library and offers a statistically principled, error-aware complement to deterministic analytic shape screening.

Supporting Information Available

Derivation of the unbiasedness and variance of the estimators; grid-convergence study; alignment and conformer-overlay experiments; software and reproducibility details (PDF).

References

(1) Hawkins, P. C. D.; Skillman, A. G.; Nicholls, A. Comparison of Shape-Matching and Docking as Virtual Screening Tools. *J. Med. Chem.* **2007**, *50*, 74–82.

(2) Kumar, A.; Zhang, K. Y. J. Advances in the Development of Shape Similarity Methods and Their Application in Drug Discovery. *Front. Chem.* **2018**, *6*, 315.

(3) Grant, J. A.; Pickup, B. T. A Gaussian Description of Molecular Shape. *J. Phys. Chem.* **1995**, *99*, 3503–3510.

(4) Grant, J. A.; Gallardo, M. A.; Pickup, B. T. A Fast Method of Molecular Shape Comparison: A Simple Application of a Gaussian Description of Molecular Shape. *J. Comput. Chem.* **1996**, *17*, 1653–1666.

(5) Yan, X.; Gu, Q.; Lu, F.; Li, J.; Xu, J. Enhancing Molecular Shape Comparison by Weighted Gaussian Functions. *J. Chem. Inf. Model.* **2013**, *53*, 1967–1978.

(6) Yan, X.; Li, J.; Gu, Q.; Xu, J. gWEGA: GPU-Accelerated WEGA for Molecular Superposition and Shape Comparison. *J. Comput. Chem.* **2014**, *35*, 1122–1130.

(7) Haque, I. S.; Pande, V. S. PAPER—Accelerating Parallel Evaluations of ROCS. *J. Comput. Chem.* **2010**, *31*, 117–132.

(8) Atwi, R.; Wang, Y.; others ROSHAMBO: Open-Source Molecular Alignment and 3D Similarity Scoring. *J. Chem. Inf. Model.* **2024**,

(9) Atwi, R.; others ROSHAMBO2: Accelerating Molecular Alignment for Large Chemical Libraries with GPU Optimization and Algorithmic Advances. *J. Chem. Inf. Model.* **2025**,

(10) Ballester, P. J.; Richards, W. G. Ultrafast Shape Recognition to Search Compound Databases for Similar Molecular Shapes. *J. Comput. Chem.* **2007**, *28*, 1711–1723.

(11) Schreyer, A. M.; Blundell, T. USRCAT: Real-Time Ultrafast Shape Recognition with Pharmacophoric Constraints. *J. Cheminform.* **2012**, *4*, 27.

(12) Jebara, T.; Kondor, R.; Howard, A. Probability Product Kernels. *J. Mach. Learn. Res.* **2004**, *5*, 819–844.

(13) Owen, A. B.; Maximov, Y.; Chertkov, M. Importance Sampling the Union of Rare Events with an Application to Power Systems Analysis. *Electron. J. Stat.* **2019**, *13*, 231–254.

(14) Lukashevich, A.; Gorchakov, V.; Vorobev, P.; Deka, D.; Maximov, Y. Importance Sampling Approach to Chance-Constrained DC Optimal Power Flow. *IEEE Trans. Control Netw. Syst.* **2023**, *11*, 928–937.

(15) Mitrovic, M.; Lukashevich, A.; Vorobev, P.; Terzija, V.; Budennyy, S.; Maximov, Y.; Deka, D. Data-Driven Stochastic AC-OPF Using Gaussian Process Regression. *Int. J. Electr. Power Energy Syst.* **2023**, *152*, 109249.

(16) Likhosherstov, V.; Maximov, Y.; Chertkov, M. Inference and Sampling of K_{33}-free Ising Models. Proceedings of the 36th International Conference on Machine Learning (ICML). 2019; pp 3963–3972.

(17) Lukashevich, A.; Bulkin, A.; Maximov, Y. A-Priori Reduction of Scenario Approximation for Automated Generation Control in High-Voltage Power Grids with Renewable Energy. *IEEE Control Syst. Lett.* **2024**, *8*, 1613–1618.

(18) Mohamed, S.; Rosca, M.; Figurnov, M.; Mnih, A. Monte Carlo Gradient Estimation in Machine Learning. *J. Mach. Learn. Res.* **2020**, *21*, 1–62.

(19) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. Int. Conf. on Learning Representations (ICLR). 2015; arXiv:1412.6980.

(20) Audibert, J.-Y.; Bubeck, S.; Munos, R. Best Arm Identification in Multi-Armed Bandits. Conf. on Learning Theory (COLT). 2010; pp 41–53.

(21) Bradbury, J.; Frostig, R.; Hawkins, P.; Johnson, M. J.; Leary, C.; Maclaurin, D.; Necula, G.; Paszke, A.; VanderPlas, J.; Wanderman-Milne, S.; Zhang, Q. JAX: Compos-

able Transformations of Python+NumPy Programs. `http://github.com/google/jax`, 2018.

(22) Riniker, S.; Landrum, G. A. Better Informed Distance Geometry: Using What We Know To Improve Conformation Generation. *J. Chem. Inf. Model.* **2015**, *55*, 2562–2574.

(23) Mysinger, M. M.; Carchia, M.; Irwin, J. J.; Shoichet, B. K. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J. Med. Chem.* **2012**, *55*, 6582–6594.

(24) Tran-Nguyen, V.-K.; Jacquemard, C.; Rognan, D. LIT-PCBA: An Unbiased Data Set for Machine Learning and Virtual Screening. *J. Chem. Inf. Model.* **2020**, *60*, 4263–4273.

(25) Truchon, J.-F.; Bayly, C. I. Evaluating Virtual Screening Methods: Good and Bad Metrics for the "Early Recognition" Problem. *J. Chem. Inf. Model.* **2007**, *47*, 488–508.