



Scatter-Limited Hybrid Monte Carlo, Deterministic Transport with Quasi-Monte Carlo Sampling

Johannes Krotz ^{1*}, Ryan G. McClarren¹

¹Department of Aerospace and Mechanical Engineering, University of Notre Dame, Notre Dame, IN

Under Review

ABSTRACT

We present a hybrid method for time-dependent particle transport that combines Monte Carlo (MC) estimation with a deterministic discrete ordinates (S_N) solve, augmented by quasi-Monte Carlo (QMC) sampling. For spatial discretizations, the MC component computes a piecewise-constant (cell-averaged) solution, while the S_N stage employs bilinear discontinuous finite elements. By hybridizing the formulation, the MC subproblem after a prescribed scatter limit becomes scattering-free, yielding a simple and efficient streaming/attenuation procedure. Between time steps, a simple scatter-free MC step is run to relabel the S_N solution as an MC solution.

A key feature of the approach is a tunable parameter N_s that controls how many material collisions are handled in the (Q)MC leg before handing off to the deterministic S_N solve; $N_s = 0$ recovers a purely uncollided MC leg, while $N_s > 0$ produces multi-scatter hybrids. QMC replaces pseudorandom draws with low-discrepancy points in the existing MC sampling maps, enabling a plug-in adoption within the standard MC code with modest, localized changes.

We observe significant accuracy and convergence rate improvements through the use of QMC and practically no additional computational cost, which are generally not seen in comparable non-hybrid solves. We believe the multi-scatter approach provides additional flexibility in terms of parallelization and the choice of deterministic solver.

Keywords: Hybrid stochastic-deterministic method, Monte Carlo, Kinetic equations Particle transport

1. INTRODUCTION

Neutron transport is a crucial phenomenon in reactor physics. Deterministic discretizations of the full phase space are accurate but costly in seven dimensions (3 space, 2 direction of flight, 1 Energy, 1 time) [1], whereas Monte Carlo (MC) handles complex geometries and continuous physics but carries $N_p^{-1/2}$ sampling error [2]. Hybrid strategies seek the best of both: regional-coupling and biasing frameworks [3, 4], and high-order/low-order (HOLO) closures [5]. A complementary line exploits *streaming-scattering splitting*, treating potentially anisotropic streaming separately from the more isotropic scattered field [6, 7, 8]. Related discontinuous Galerkin (DG)–DG hybrids have demonstrated multigroup coarsening for collided components [9].

Contributions. This paper introduces a time-dependent MC–DG hybrid that: (i) advances particles with MC *up to a fixed cap of N_s material scatterings* per step (the MC leg is not scatter-free for $N_s > 0$); (ii) transfers particles exceeding N_s to a collided equation solved by discrete ordinates with a *discontinuous Galerkin (DG)* spatial discretization; and (iii) applies a stepwise remap that resamples from the DG solution back into MC to prevent exponential depletion of low-scatter particles. With an appropriate

*jkrotz@nd.edu

DG scheme, the method retains the diffusion limit [10]. Numerical experiments show that this hybrid approach sees improved convergence rates by a simple replacement of pseudo-random numbers with quasi-random numbers. The bounded multi-scatter MC mechanism is new and developed in the next section. We focus on single-group MC–DG.

2. BASICS OF THE QMC–HYBRID METHOD

2.1. Transport equation

Let $X \subset \mathbb{R}^3$ be a spatial domain with Lipschitz boundary and let S^2 be the unit sphere in \mathbb{R}^3 . Let $\Psi = \Psi(x, \Omega, t)$ be the angular flux depending on the position $x = (x, y, z) \in X$, the direction of flight $\Omega \in S^2$ and time $t > 0$. We assume that Ψ is governed by the linear transport equation

$$\frac{1}{c} \partial_t \Psi + \Omega \cdot \nabla_x \Psi + \sigma_a \Psi = \sigma_s \mathcal{S}[\Psi] + Q, \quad x \in X, \quad \Omega \in S^2, \quad t > 0, \quad (1)$$

where $\sigma_s = \sigma_s(x)$ and $\sigma_a = \sigma_a(x)$ are the scattering and absorption cross-sections of the material, respectively; $Q = Q(x, \Omega, t)$ is a known particle source; and \mathcal{S} denotes the scattering kernel, which throughout this paper will be:

$$\mathcal{S}[\Psi] = \frac{\langle \Psi \rangle}{4\pi} - \Psi = \frac{1}{4\pi} \int_{S^2} \Psi d\Omega - \Psi. \quad (2)$$

Note that isotropic scattering is not a restriction of the method. Indeed, this hybrid approach could be readily extended to that scattering in future work.

The constant $c > 0$ is the particle speed; we assume that $c = 1$ for the remainder of this work. The transport equation (1) is equipped with initial conditions

$$\Psi(x, \Omega, 0) = \tilde{\Psi}(x, \Omega), \quad x \in X, \quad \Omega \in S^2, \quad (3)$$

and boundary condition

$$\Psi(x, \Omega, t) = G(x, \Omega, t) \quad x \in \partial X, \quad \Omega \cdot n(x) < 0, \quad (4)$$

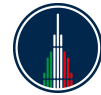
where Ψ_0 and G are known and $n(x)$ is the unit outward normal at $x \in \partial X$.

2.2. The hybrid method

The hybrid method is based on a collision source splitting [11]. Let $\Psi = \sum_{n=0}^{\infty} \Psi_n$, where the n -collision flux $\{\Psi_n\}_{n=0}^{\infty}$ satisfy the following system of equations

$$\partial_t \Psi_n + \Omega \cdot \nabla_x \Psi_n + \sigma_a \Psi_n = Q_n - \sigma_s \Psi_n, \quad n \in \mathbb{N}_0, \quad (5.n)$$

with source terms $Q_0 = Q$ and $Q_n = \frac{\sigma_s}{4\pi} \langle \Psi_{n-1} \rangle$ for all $n > 0$. For simplicity we assume that $\Psi_0(x, \Omega, 0) = \tilde{\Psi}(x, \Omega)$ and $\Psi_0(x, \Omega, t) = G(x, \Omega, t)$ on $x \in \partial X$ for incoming directions. Thus the initial and boundary data for $n > 0$ end up being 0. This split divides particles into sets with densities Ψ_n , in which each particle has currently undergone n collisions since its birth at the source Q or entry into the domain. This further splits the scattering operator $\sigma_s \mathcal{S}[\Psi_n]$ into an absorption term $-\sigma_s \Psi_n$ and a reemission term $\frac{\langle \sigma_s \Psi_n \rangle}{4\pi}$. The absorption parts shows up as a sink in the equation for Ψ_n , which is interpreted as a particle with n collisions disappearing as particle with n collisions, when it scatters an additional time, effectively being



absorbed. While the isotropic reemission appears in the equation for Ψ_{n+1} indicating that an n -scattered particle is remitted with $n + 1$ collision after an additional scattering event.

Due to the linearity of (1), the splitting in (5.n) is exact; that is, if all Ψ_n solve their respective (5.n), then $\sum_{n=0}^{\infty} \Psi_n$ solves (1). A hybrid method will solve equations (5.n) in two batches by splitting $\Psi = \sum_{n=0}^{\infty} \Psi_n = \Psi_0 + \sum_{n=1}^{\infty} \Psi_n =: \Psi_u + \Psi_c$, into the densities of uncollided and collided particles. We build on this idea by splitting this some along a different $N_s \neq 0$ into

$\Psi = \sum_{n=0}^{N_s-1} \Psi_n + \Psi_{N_s} + \sum_{n=N_s+1}^{\infty} \Psi_n =: \Psi_{\lceil N_s \rceil} + \Psi_{N_s} + \Psi_{\lfloor N_s+1 \rfloor}$ with the *sub- N_s collision flux* $\Psi_{\lceil N_s \rceil}$, the *N_s -collision flux* and the *super- N_s -collision flux* $\Psi_{\lfloor N_s+1 \rfloor}$ solving

$$\begin{aligned} \partial_t \left(\Psi_{\lceil N_s \rceil} + \Psi_{N_s} \right) + \Omega \cdot \nabla_x \left(\Psi_{\lceil N_s \rceil} + \Psi_{N_s} \right) + \sigma_a \Psi_{\lceil N_s \rceil} + \sigma_t \Psi_{N_s} &= Q + \sigma_s \mathcal{S} \left[\Psi_{\lceil N_s \rceil} \right] \quad (6. \lceil N_s \rceil) \\ \partial_t \Psi_{\lfloor N_s+1 \rfloor} + \Omega \cdot \nabla_x \Psi_{\lfloor N_s+1 \rfloor} + \sigma_t \Psi_{\lfloor N_s+1 \rfloor} &= \sigma_s \mathcal{S} \left[\Psi_{\lfloor N_s+1 \rfloor} \right] + \frac{\sigma_s}{4\pi} \langle \Psi_{N_s} \rangle \quad (6. \lfloor N_s+1 \rfloor) \end{aligned}$$

This is a split into particles that scattered at most N_s times and particles that scattered at least $N_s + 1$ times. For $N_s = 0$ the sum in $\Psi_{\lceil N_s \rceil}$ is empty and previous hybrid methods are recovered. Note that a solution to (6. $\lceil N_s \rceil$) consists of the pair $(\Psi_{\lceil N_s \rceil}, \Psi_{N_s})$. Further note that for $N_s > 0$ this, on its own, is not a well-posed problem and allows for infinitely many solutions. Throughout this manuscript this will be resolved by conditioning $\{\Psi_n\}_0^{N_s}$ to solve (5.n). This determines all Ψ_n and thus $\Psi_{\lceil N_s \rceil} \Psi_{N_s}$ uniquely. A short discussion of this is included in Appendix A.

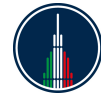
In practice, (6. $\lceil N_s \rceil$) and (6. $\lfloor N_s+1 \rfloor$) are solved at different resolutions or even with different methods. Typically (6. $\lceil N_s \rceil$) is solved with a method that has high resolution in angle, and because (6. $\lceil N_s \rceil$) has no ($N_s = 0$), or lower ($N_s > 0$ but finite) coupling in angle, it is easier to solve than (1) and also easier to solve in parallel. Meanwhile (6. $\lfloor N_s+1 \rfloor$) inherits the full angular coupling in (1), but typically requires less angular resolution. Within this work we will assume that (6. $\lceil N_s \rceil$) is solved using a (Q)MC-simulation, described in the next section, and (6. $\lfloor N_s+1 \rfloor$) is solved using an S_N -solver, for example described in [6].

Previous hybrid solvers for this system have, to our knowledge, only covered the case of $N_s = 0$ corresponding to a split into uncollided particles and particles having undergone at least 1 scattering event. In this special case $\Psi_{\lceil N_s \rceil} \equiv 0$, simplifying (6. $\lceil N_s \rceil$). In the version discussed going forward, we will present a hybrid that can solve these equations for arbitrary N_s . This allows for a tunable division of workloads between the different methods used for each part.

Assuming (6. $\lceil N_s \rceil$) is solved at high accuracy, while (6. $\lfloor N_s+1 \rfloor$) is done cheaply, a majority of the workload should be carried out in the former of the two to retain high accuracy overall. However, in the non-trivial case with $\sigma_s \neq 0$ particles scatter at an exponential rate, thus particle densities are, over time, transferred out of Ψ_n into Ψ_{n+1} for every n . Thus particle densities will be transferred into the N_s -tail flux at an exponential rate, thus making the accuracy at which (6. $\lfloor N_s+1 \rfloor$) is solved the driving factor in the overall accuracy. We briefly give bounds on this transfer rate in Appendix B

This effect can be mitigated by abusing the autonomous nature of the equations and periodically relabelling the high scatter densities as uncollided at every time step. This leads to the following algorithm to propagate a solution $\Psi[t_n](x, \Omega) = \Psi(x, \Omega, t_n)$ along the time step $[t_n, t_{n+1}]$ to $\Psi[t_{n+1}]$:

1. With initial data $\Psi_{\lceil N_s \rceil}^{MC}[t_n] = \Psi[t_n]$ and $\Psi_{N_s}^{MC} \equiv 0$ solve (6. $\lceil N_s \rceil$) for solutions $\Psi_{\lceil N_s \rceil}^{MC}[t_{n+1}]$ and $\Psi_{N_s}^{MC}[t_{n+1}]$ via the scatter-limited MC algorithm with N_s scatters described in the next section. Apply source $Q[t]$ and boundary data $G[t]$ in this step.
2. Solve (6. $\lfloor N_s+1 \rfloor$) for an S_N -solution $\Psi^{SN}[t_{n+1}]$ with zero initial and boundary data.



3. (Remap) With zero allowable scatters in the MC–algorithm solve

$$\partial_t \Psi_R + \Omega \cdot \nabla_x \Psi_R + \sigma_t \Psi_R = \left(Q + \sigma_s \langle \Psi_{[N_s]} \rangle^{MC} \right) + \sigma_s \langle \Psi_{[N_s]} \rangle^{MC} + \frac{\sigma_s}{4\pi} \langle \Psi_{N_s} \rangle^{SN} \quad (6.R)$$

for $\Psi_R^{MC}[t_{n+1}]$. Use initial data $\Psi_R[t_n] = \Psi[t_n]$ and apply boundary data $G[t]$.

4. Set $\Psi[t_{n+1}] = \Psi_R^{MC}[t_{n+1}]$.

Note that Ψ^{MC} will generally be represented by a collection of computer particles distributed proportional to Ψ^{MC} , $\langle \Psi \rangle^{MC/SN}$ are MC and S_N estimators of $\langle \Psi \rangle$ respectively, in our case an implicit capture track length estimator in the MC case and Gauss-Legendre integration of the ordinates for S_N .

Further note that previous hybrids replaced steps 3 and 4 with

3'. With zero allowable scatters in the MC–algorithm solve

$$\partial_t \Psi_R + \Omega \cdot \nabla_x \Psi_R + \sigma_t \Psi_R = \sigma_s \langle \Psi_{[N_s]} \rangle^{MC} + \frac{\sigma_s}{4\pi} \langle \Psi_{N_s} \rangle^{SN} \quad (6.R')$$

for $\Psi_R^{MC}[t_{n+1}]$. Boundary and initial data in this case are zero.

4'. Set $\Psi[t_{n+1}] = \Psi_{[N_s]}^{MC}[t_{n+1}] + \Psi_{N_s}^{MC}[t_{n+1}] + \Psi_R^{MC}[t_{n+1}]$.

These approaches are, in theory, equivalent due to the linearity of the equations involved. However, we found that the version with 3' and 4' does not benefit from QMC. Further, step 3' requires the addition of new computer particles, whereas 3 can be solved by resetting the particles from 1, thus allowing for much easier control of overall particle numbers. In practice 3 can be run by resetting particles that were added through the source terms, while particles from initial data and boundary terms can just be copied from 1 without additional computation.

2.3. Scatter-limited Monte Carlo (MC) Algorithm

The algorithm described in the previous section requires an (Q)MC solver for $(6.[N_s])$ that (a) will scatter at most N_s times and (b) tracks $\langle \Psi_{[N_s]} \rangle$ and $\langle \Psi_{N_s} \rangle$ separately. First we will briefly discuss the unlimited scatters case ($N_s = \infty$ and $\Psi_{N_s} = \Psi_\infty \equiv 0$), as well as the zero scatter case ($N_s = 0$ and $\Psi_{[N_s]} = \Psi_{[0]} \equiv 0$). In all cases we use implicit capture. Thus we start by sampling particles with birth location and time proportional to $q(x, \Omega, t) = Q(x, \Omega, t) / \int_{X \times S^2 \times I} Q dx d\Omega dt$ with isotropic Ω and assign weights w_0 so that $\sum_p w_0 = \int Q$. Particle weights will decrease exponentially along their trajectory.

Unlimited (reference) case. Particles stream on straight lines until they scatter; along a segment of length ℓ , apply implicit capture $w \leftarrow w \exp\left(-\int_0^\ell \sigma_a ds\right)$. Candidate scatters are drawn at exponential rate σ_s ; upon a collision, resample Ω isotropically and continue. Terminate on exit, end of step, or $w < w_{\min}$. Track the angle-average $\langle \Psi \rangle = (4\pi)^{-1} \int_{S^2} \Psi d\Omega$ via a track-length estimator in space–time bins.

Non-scattering variant. Particles never change direction, but lose weight with rate $\sigma_t = \sigma_a + \sigma_s$, i.e., $w \leftarrow w \exp\left(-\int \sigma_t ds\right)$. The resulting $\langle \Psi \rangle$ along trajectories is supplied to another solver to account for higher-order scattering.

N_s -scatter-limited variant. At birth or the start of the time step each particle is assigned a scatter counter n . The simulation of a particle changes based on its current scatter count.



- *Pre-limit* ($n < N_s$): between collisions, use implicit capture with σ_a and sample next scatter at rate σ_s ; after each scatter, set $n \leftarrow n + 1$ and resample Ω .
- *Post-limit* ($n = N_s$): no further collisions; stream with attenuation σ_t only.

Maintain two tallies: $\langle \Psi \rangle_{\text{pre}}$ for segments with $n \leq N_s - 1$ and $\langle \Psi \rangle_{\text{post}}$ for the $n = N_s$ leg. Use $\langle \Psi \rangle_{\text{post}}$ to form the exactly- N_s collided source, e.g., $Q_{N_s+1} = (\sigma_s/4\pi) \langle \Psi_{N_s} \rangle$, and iterate to recover (N_s+1) -and-higher collisions. For $N_s = 0$, this reduces to the standard uncollided/collided split.

Pseudocode (one time step, inline). Initialize $(x, \Omega, t) \sim q$, $w \leftarrow w_0$, $n \leftarrow 0$. While alive: if $n < N_s$, sample a candidate collision at rate σ_s , stream to the nearest of boundary/step end/collision, tally into $\langle \Psi \rangle_{\text{pre}}$, attenuate w with σ_a ; if a collision occurs, resample Ω , set $n \leftarrow n + 1$. Else ($n = N_s$): stream to termination without collisions, tally into $\langle \Psi \rangle_{\text{post}}$, attenuate with σ_t . Terminate on exit, step end, or $w < w_{\text{min}}$. Finally, pass $\langle \Psi \rangle_{\text{post}}$ to the collided solver and iterate as needed.

2.4. Quasi Monte Carlo (QMC)

Quasi-Monte Carlo (QMC) methods replace pseudorandom draws on the unit cube with low-discrepancy points to reduce variance for a given sample size. In simulation, these unit-cube points are typically transformed to the required physical inputs via inverse-CDF maps (or equivalent measure-preserving transforms). In this work we use a Halton sequence because it is simple, extensible in dimension, and performs robustly for the moderate dimensions relevant here.

For transport simulation, as discussed here, a natural QMC approach is to first sample the Poisson-distributed number of scattering events to fix the stochastic dimension, and then use the low-discrepancy points to sample the scattering locations/directions. However, in non-homogeneous materials, the event rate varies along the path, so the number of scatter events is path-dependent the required transform is not known a priori. This makes a fixed-dimension QMC mapping cumbersome and usually impractical, eroding the advantage of low-discrepancy sampling in that setting. In our hybrid approach, we see improved convergence rates by simply replacing the pseudo-random numbers of the MC algorithm with the quasi-random numbers given by a Halton sequence. We believe this to be due to the resampling process being scatter-free, effectively limiting the required dimension of random numbers to the degrees of freedom defining such a straight-line trajectory.

3. NUMERICAL EXAMPLES

As numerical examples we include Reed's problem [12], a 1D problem with heterogenous materials and commonly used as a test case as well as a 2D problem based on the Kobayashi Dogleg benchmark [13]. Geometries and material properties for either problem are included in the Appendix C. Both problems are run as equilibrium problems, i.e. with a single infinite time step. While the algorithm is written for and works on explicitly time dependent problems as well, we defer to future work for numerical examples. We restrict our analysis to cases with an S_4 -solver for $\Psi_{[N_s]}$. The solver specifics are outlined in [6]. In that work we also demonstrated that in the $N_s = 0$ case the order of the discrete ordinates solver does affect the overall accuracy marginally at best. Additional tests for $N_s > 0$, not discussed here, demonstrate that this remains true here as well. All runs were performed sequentially on the same local machine using code built on top of MCDC, developed by the Center for Exascale Monte-Carlo Neutron Transport (CEMeNT) [14]. Results depicted in figure 1 focus on the convergence in L^2 and run time as a function of the number of particles, number of scatter events per particle N_s and based on whether QMC or regular MC was used. The last column in these figures depicts the L^2 -errors as function of the runtime, to analyze overall computational efficiency: more efficient runs appear either lower (more accurate), further to the left (faster) than less efficient runs, or both. Spatial resolution was kept constant throughout with $N_x = 80$ for Reed's problem and $N_x \times N_y = 30 \times 50$ for the Dogleg benchmark. N_s was varied between $\{0, 5, 10, 20, \infty\}$, with $N_s = \infty$ corresponding to a pure MC simulation. Relabeling is

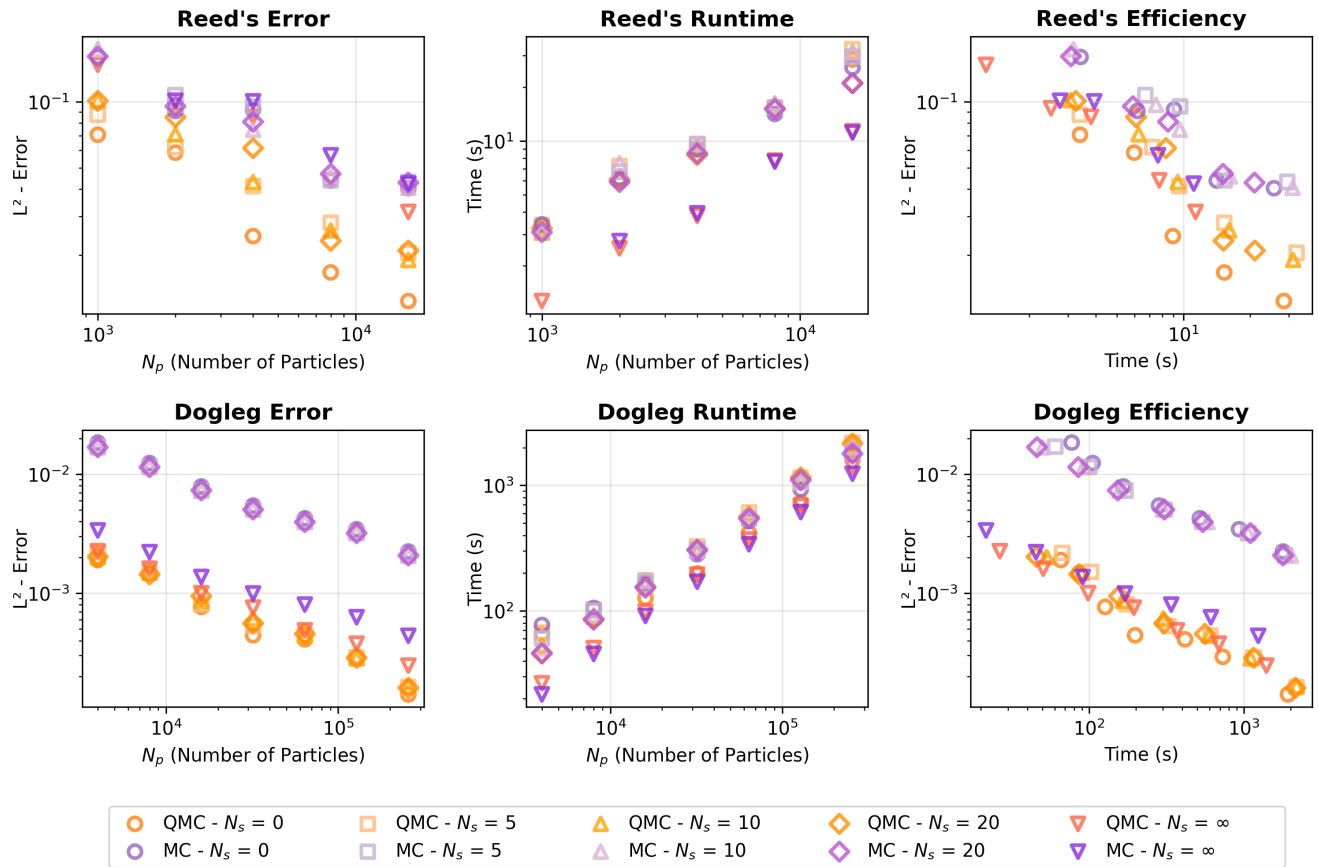


Figure 1. Summary of numerical experiments for Reed's problem (top row) and the Dogleg problem (bottom row). Depicted are the L^2 -error as depending on the particle count N_p (left column), the runtime as depending on N_p (middle column) and L^2 -error depending on the runtime (right column) as a measure of efficiency. Runs that appear closer to the bottom left provide higher accuracy per runtime. We compare QMC runs (orange) to MC runs (purple) for various values of the scatter limit $N_s \in \{0, 5, 10, 20\}$ as well as $N_s = \infty$ (non-hybrid/regular (Q)MC run).

unnecessary in this case and therefore skipped.

Significant key observations in these figures are the following: The hybrid methods, independently of N_s , see significant improvements in accuracy by using QMC at no additional cost in terms of runtime. This effect is much less pronounced for the $N_s = \infty$, i.e. the non-hybrid case. Hybrid runs that use QMC over MC appear to be more accurate per particle than non-hybrid runs. This is solidified by improved convergence rates of QMC runs for the hybrid method, as outlined in table I. This still holds for regular MC in the 1D, but not in the 2D example. It is not clear whether this is a question of dimensionality or problem dependent. Non-hybrid runs are, for the same number of particles, faster. This is expected as no relabeling is happening. For QMC runs the additional accuracy of hybrid runs makes up for the longer runtime, leading to QMC hybrid runs being comparably or more efficient than non-hybrid runs. For hybrid runs the effect of different N_s appears to have little effect on the run time. This is likely due to the fact that higher N_s results in what seems to be proportional reduction in the number of iterations needed for the S_N -solver, also shown in table I.

Table I. Comparison of number of S_N iterations and convergence rate as function of allowable number of scatter events N_s for Reed's problem(left) and the Dogleg problem (right). S_N iterations are averaged over runs with various N_p , but N_p dependence appears negligible. Convergence rate α is determined via least-square fit of cN_p^α to the L^2 -distance between approximate and reference solutions. We distinguish α_{MC} and α_{QMC} based on the random numbers used.

N_s	Reed's problem			Dogleg		
	S_N -Iterations	α_{QMC}	α_{MC}	S_N -Iterations	α_{QMC}	α_{MC}
0	31	-0.69	-0.5	7	-0.61	-0.49
5	26	-0.64	-0.51	4	-0.61	-0.49
10	21.4	-0.63	-0.53	1	-0.61	-0.49
20	12.2	-0.64	-0.49	1	-0.61	-0.49
∞	0	-0.56	-0.46	0	-0.53	-.47

4. CONCLUSIONS

We showed that with the addition of a relatively cheap and standard deterministic solver and a relabeling step, standard MC algorithms for particle transport can reap the benefits of that come with the use of quasi-random numbers with only minor, localized changes. Namely (i) replacing pseudorandom draws by low-discrepancy points without altering the solver's existing inverse-CDF/transport maps, (ii) inserting a simple scatter counter with a post-limit non-scattering leg, and (iii) tallying $\langle \Psi \rangle$ before/after the N_s -th scatter—transport solvers. This gives a plug-in pathway to QMC *without* substantial changes to the current sampling procedure. This keeps code disruption low while unlocking QMC's variance-reduction benefits inside otherwise conventional workflows.

We also introduced *multi-scatter hybrids* ($N_s > 0$), which solve a predetermined number of collisions with (Q)MC and delegate the remainder to a deterministic “clean-up” solve for highly scattered particles. Although the direct impact of multi-scattering on accuracy and wall time was marginal in our tests, this avenue is still compelling: We only reported runtimes in a sequential setting. However, multi-scatter QMC is far easier to parallelize than the coupled clean-up stage, promising speedups in non-sequential (e.g., many-core or distributed) runs.

We observe, not unexpectedly, a direct trade-off between the number of iterations used in the ‘clean up’ and the number of allowable scatters in the (Q)MC-stage.

A practical tuning knob emerges from this tradeoff between the number of MC-handled scatters and the iterations required by the clean-up solver: larger N_s reduces angular-coupling burden downstream, enabling cheaper deterministic settings in high-scattering regimes. This division of labor can be automated to target either fixed accuracy at minimum time or maximum accuracy under a time budget. One can envision a hybrid with a fully parallel (Q)MC leg and relatively high N_s , combined with a deterministic solver as cheap as a simple diffusion approach, for minimal runtimes for example. Other examples could include larger, N_s -dependent time steps, as with increased N_s fewer particles enter the deterministic regime.

Finally, throughout this manuscript we chose to resolve the non-uniqueness of equation (6. $\lceil N_s \rceil$) by conditioning Ψ_n to solve (5.n) for all $n \leq N_s$ for a fixed N_s . This is an intuitive choice, but far from the only one, leaving tremendous space for future design choices. As a simple next step N_s can be made to depend on space, cross sections, time, or other indicators (e.g., optical thickness, local stiffness, or estimated variance), opening a broad path for adaptive strategies and future exploration of problem-dependent

multi-scatter schedules and plug-and-play QMC integrations.

ACKNOWLEDGEMENTS

This work was supported by the Center for Exascale Monte-Carlo Neutron Transport (CEMeNT), a PSAAP-III project funded by the Department of Energy under Grant DE-NA003967.



REFERENCES

- [1] E. E. Lewis and W. F. Miller. *Computational methods of neutron transport*. John Wiley and Sons, Inc., New York, NY (1984).
- [2] J. Spanier and E. M. Gelbard. *Monte Carlo principles and neutron transport problems*. Courier Corporation (2008).
- [3] S. W. Mosher and F. Rahnema. “The incident flux response expansion method for heterogeneous coarse mesh transport problems.” *Transport Theory and Statistical Physics*, **volume 35**(1-2), pp. 55–86 (2006).
- [4] J. C. Wagner, D. E. Peplow, and S. W. Mosher. “FW-CADIS method for global and regional variance reduction of Monte Carlo radiation transport calculations.” *Nuclear Science and Engineering*, **volume 176**(1), pp. 37–57 (2014).
- [5] H. Park, D. Knoll, R. Rauenzahn, A. Wollaber, and R. Lowrie. “Moment-based acceleration of Monte Carlo solution for multifrequency thermal radiative transfer problems.” *Journal of Computational and Theoretical Transport*, **volume 43**(1-7), pp. 314–335 (2014).
- [6] J. Krotz, C. D. Hauck, and R. G. McClarren. “A hybrid Monte Carlo, discontinuous Galerkin method for linear kinetic transport equations.” *Journal of Computational Physics*, **volume 514**, p. 113253 (2024).
- [7] C. D. Hauck and R. G. McClarren. “A collision-based hybrid method for time-dependent, linear, kinetic transport equations.” *Multiscale Modeling & Simulation*, **volume 11**(4), pp. 1197–1227 (2013).
- [8] M. M. Crockatt, A. J. Christlieb, C. K. Garrett, and C. D. Hauck. “Hybrid methods for radiation transport using diagonally implicit runge–kutta and space–time discontinuous galerkin time integration.” *Journal of Computational Physics*, **volume 376**, pp. 455–477 (2019).
- [9] B. Whewell, R. G. McClarren, C. D. Hauck, and M. Shin. “Multigroup Neutron Transport Using a Collision-Based Hybrid Method.” *Nuclear science and engineering*, pp. 1–20 (2023).
- [10] E. W. Larsen, J. E. Morel, and W. F. Miller Jr. “Asymptotic solutions of numerical transport problems in optically thick, diffusive regimes.” *Journal of Computational Physics*, **volume 69**(2), pp. 283–324 (1987).
- [11] R. E. Alcouffe. “A first collision source method for coupling Monte Carlo and discrete ordinates for localized source problems.” In *Monte-Carlo Methods and Applications in Neutronics, Photonics and Statistical Physics: Proceedings of the Joint Los Alamos National Laboratory-Commissariat à l’Energie Atomique Meeting Held at Cadarache Castle, Provence, France April 22–26, 1985*, pp. 352–366. Springer (2006).
- [12] W. H. Reed. “New difference schemes for the neutron transport equation.” *Nuclear Science and Engineering*, **volume 46**(2), pp. 309–314 (1971).
- [13] K. Kobayashi, N. Sugimura, and Y. Nagaya. “3-D Radiation Transport Benchmark Problems and Results for Simple Geometries with Void Regions.” Technical report, OECD Nuclear Energy Agency, Nuclear Science Committee (2000).
- [14] J. P. Morgan, I. Variansyah, S. L. Pasmann, K. B. Clements, B. Cuneo, A. Mote, C. Goodman, C. Shaw, J. Northrop, R. Pankaj, E. Lame, B. Whewell, R. G. McClarren, T. S. Palmer, L. Chen, D. Y. Anistratov, C. T. Kelley, C. J. Palmer, and K. E. Niemeyer. “Monte Carlo / Dynamic Code (MC/DC): An accelerated Python package for fully transient neutron transport and rapid methods development.” *Journal of Open Source Software*, **volume 9**(96), p. 6415 (2024). URL <https://joss.theoj.org/papers/10.21105/joss.06415>.

APPENDIX A. UNIQUENESS AND ALTERNATIVE CONSTRAINTS

As a standalone equation, equation (6. $\lceil N_s \rceil$) can be rewritten as

$$\partial_t(u + v) + \mathbf{\Omega} \cdot \nabla_x(u + v) + \lambda_u u + \lambda_v v = \lambda_s \mathcal{S}[u] + s, \quad (7)$$

with material parameters λ_u, λ_v and λ_s , as well as source s and appropriate boundary and initial data for u and v . A solution to this equation is a tuple of functions (u, v) that jointly fulfill the imposed conditions. Such a tuple is not determined uniquely by (7).

For example, if we impose $v \equiv 0$ the, then unique, solution can be found by setting u to be a solution to (1) and $v \equiv 0$. If, on the other hand, $u \equiv 0$ was imposed v solves (1) with the replacements $0 \rightarrow \sigma_s$ and $\sigma_t \rightarrow \sigma_a$. (We are assuming compatible boundary conditions and initial data with respect to the imposed constraints.) All possible solutions to (7) are pointwise convex combinations of these two extreme cases. To enforce uniqueness an additional constraint $\mathbf{Law}(u, v)$ on the relation of u, v is required. Examples we have already seen are

$$\mathbf{Law}(u, v) : \{u \equiv 0\} \quad (8)$$

$$\mathbf{Law}(u, v) : \{v \equiv 0\} \quad (9)$$

$$\mathbf{Law}(u, v) : \left\{ u = \sum_{n=0}^{N_s-1} \Psi_n, v = \Psi_{N_s}, \text{ s.t. } \Psi_n \text{ solve (5.n)} \right\} \quad (10)$$

Many versions of valid $\mathbf{Law}(u, v)$ can be found through small changes to the scatter-limited MC-algorithm. Examples include, but are not limited to, making N_s spatially or material dependent. We posit that the freedom in choosing a specific relation $\mathbf{Law}(u, v)$ to force uniqueness should be seen as a feature rather than a bug.

APPENDIX B. EXPONENTIAL ENTRY INTO THE HIGH-SCATTER REGIME

In the main body of the manuscript we claim that particles will transition to the high scatter regime (more than N_s scatters) at an exponential rate, if $\sigma_s \neq 0$, necessitating relabeling to avoid the deterministic solver determining the overall accuracy. This section intends to provide brief proof of this claim. Assume first that the scattering rate is constant (per cell), $\sigma_s > 0$. Then successive inter-scatter times are i.i.d. $\text{Exp}(\sigma_s)$. If a particle has already scattered $N_s - n$ times, the additional time to accrue $n+1$ scatters is

$$T_{n+1} \sim \text{Erlang}(n+1, \sigma_s).$$

In particular, the transition time from source is $T_{N_s+1} \sim \text{Erlang}(N_s+1, \sigma_s)$ with exact tail

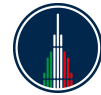
$$\mathbb{P}(T_{N_s+1} > \Delta t) = e^{-\sigma_s \Delta t} \sum_{j=0}^{N_s} \frac{(\sigma_s \Delta t)^j}{j!}, \quad \Delta t > 0.$$

With the monotonicity $\mathbb{P}(T_{n+1} > \Delta t) > \mathbb{P}(T_n > \Delta t)$, this immediately yields two-sided, explicit bounds:

$$e^{-\sigma_s \Delta t} \leq \mathbb{P}(T_0 > \Delta t) \leq \dots \leq \mathbb{P}(T_{N_s+1} > \Delta t) \leq e^{-\mu \Delta t} \left(1 - \frac{\mu}{\sigma_s}\right)^{-(N_s+1)} \quad \text{for any } \mu \in (0, \sigma_s).$$

Hence $\mathbb{P}(T_{n+1} > \Delta t) = \mathcal{O}(e^{-\mu \Delta t})$ for any fixed $\mu \in (0, \sigma_s)$: particles enter the $> N_s$ scatter regime at an *exponential* rate. This constant- σ_s assumption is sufficient for per-cell analysis. If the scattering cross sections varies but is bounded

$$0 < \sigma_s^{\min} \leq \sigma_s(\cdot) \leq \sigma_s^{\max},$$



then by coupling with homogeneous Poisson processes one has the stochastic sandwich $T_n^{(\sigma_s^{\max})} \preceq_{\text{st}} T_{n+1} \preceq_{\text{st}} T_{n+1}^{(\sigma_s^{\min})}$. Consequently

$$e^{-\sigma_s^{\max} \Delta t} \leq \mathbb{P}(T_{n+1} > \Delta t) \leq e^{-\mu \Delta t} \left(1 - \frac{\mu}{\sigma_s^{\min}}\right)^{-(N_s+1)} \quad \text{for any } \mu \in (0, \sigma_s^{\min}).$$

Thus the exponential transition rate persists under $0 < \sigma_s^{\min} \leq \sigma_s(\cdot) \leq \sigma_s^{\max}$.

APPENDIX C. GEOMETRY AND SELECT RUNS OF NUMERICAL PROBLEMS

In this section we will provide the material properties and geometries used in our numerical examples, so that a reader could reproduce these runs or use the same set ups as benchmarks for different methods. We provide outputs from some example runs as well.

C.1. Reed's problem

Reeds' problem is a 1D benchmark problem pulled from [12]. It is made up of 4 slabs of different materials reflected at $x = 0$. We removed the reflective boundary and replaced it via domain duplication. The layout simulated, including values for σ_a, σ_s and Q are depicted in figure 2. Example runs for a few different particle counts N_p and different methods (QMC, MC and non-hybrid) are depicted in figure 3.

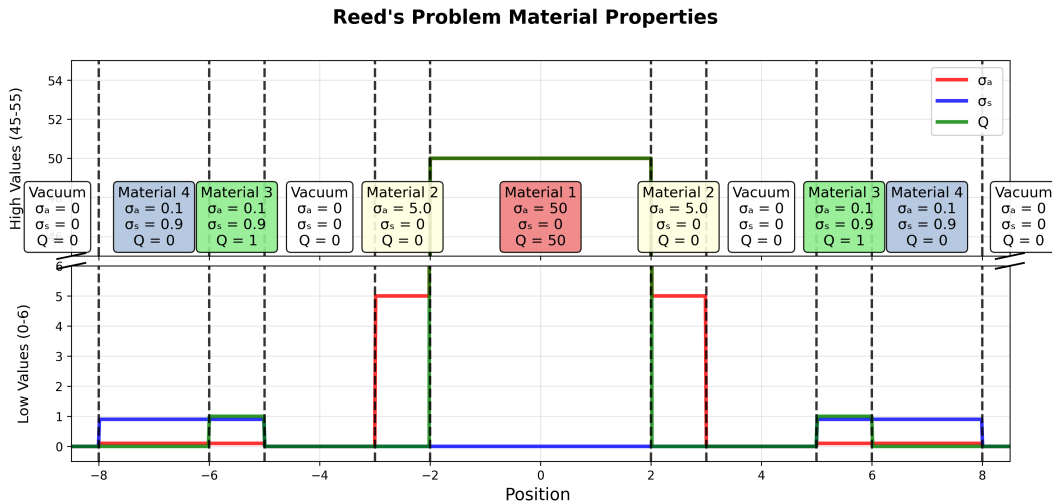


Figure 2. Material parameters and geometry for Reed's problem

C.2. Dogleg problem

The Dogleg problem is a 2D problem inspired by the 3D-Dogleg problem originally suggested in [13]. It is a simple near-vacuum duct winding through an absorbing and scattering material with a source at the start of the duct. We turned the 3D problem into a 2D problem by infinitely extending the duct and source along the z -direction. The schematic layout including values for σ_a, σ_s and Q throughout the domain is depicted in figure 4. Example runs for a few different particle counts N_p and different methods (QMC, MC and non-hybrid) are depicted in figure 5.

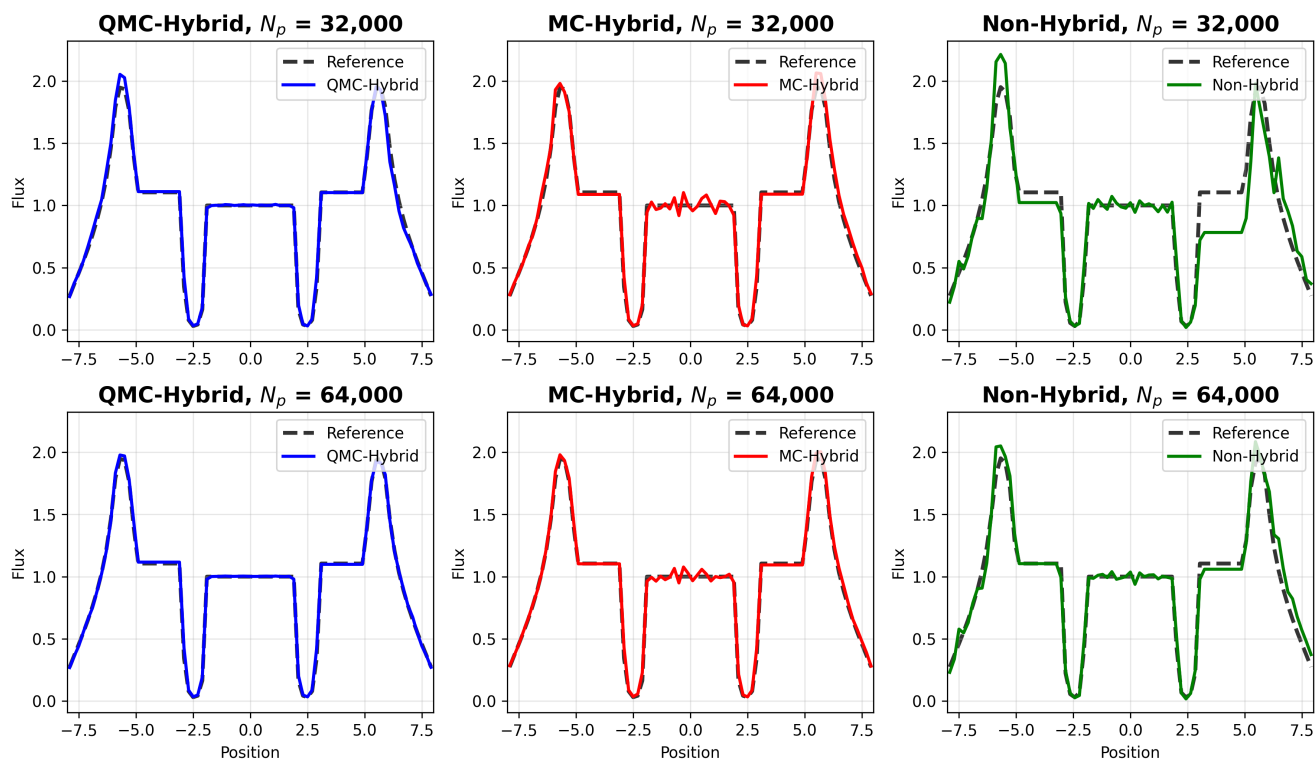


Figure 3. Sample outputs of $\langle \Psi \rangle$ for Reed's problem with $N_p = 32000$ (top row) and $N_p = 64000$ (bottom row). We compare results to a reference (dashed) for QMC-hybrid runs(left), MC-hybrid runs(middle) and non-hybrid runs(right). The scattering limit N_s is set to zero for hybrid runs and ∞ for non-hybrid runs.

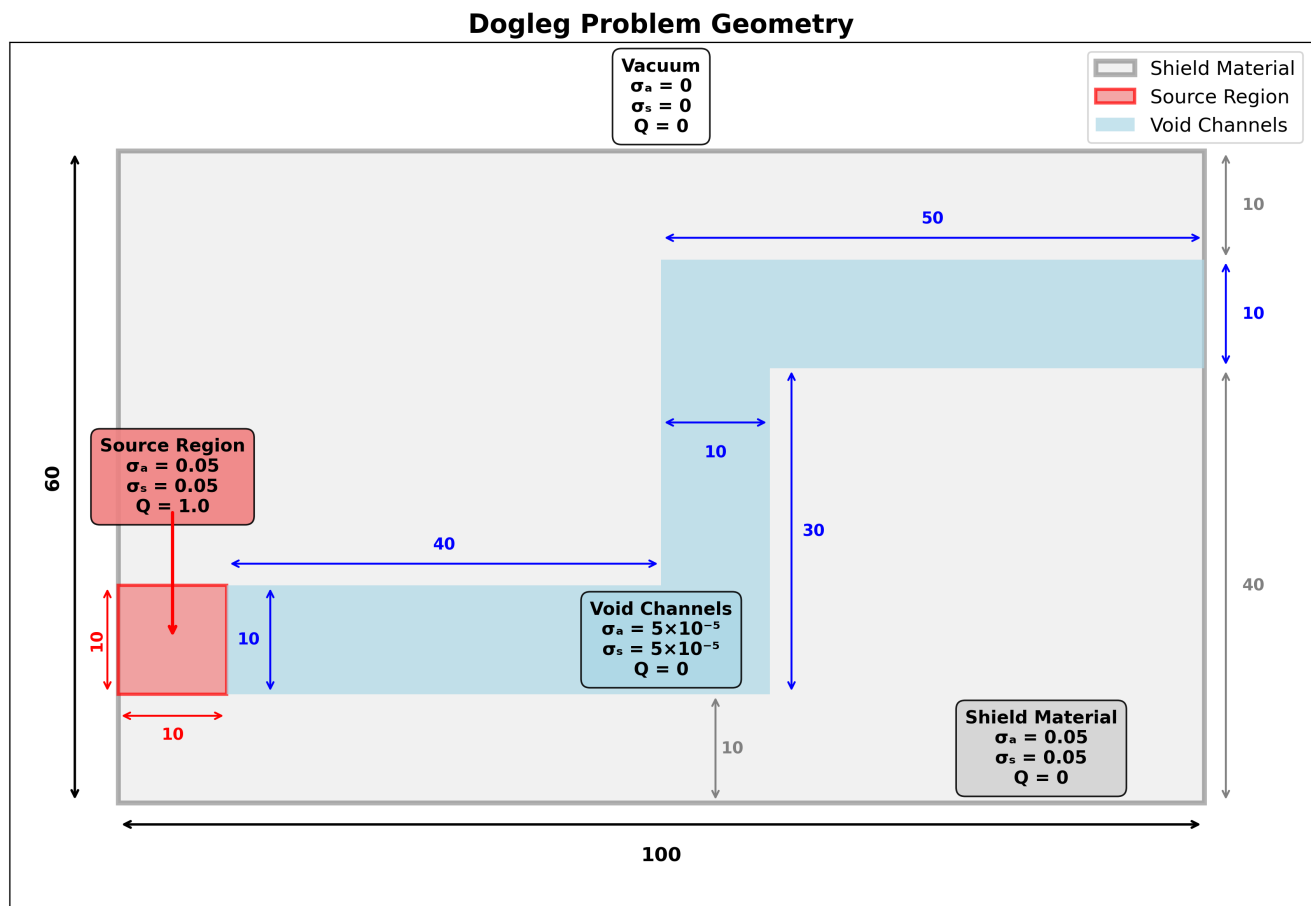


Figure 4. Material parameters and geometry for the Dogleg problem (rotated by 90°)

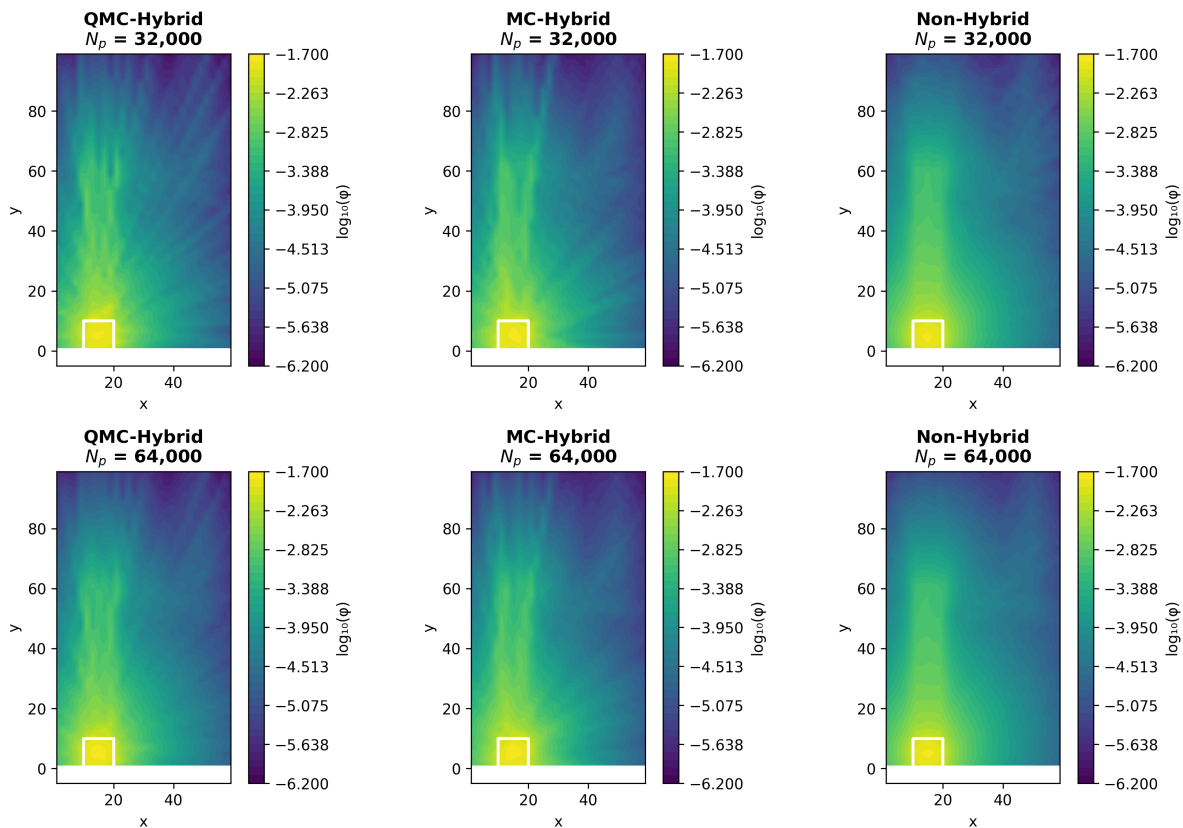


Figure 5. Sample outputs of $\langle \Psi \rangle$ for the Dogleg problem with $N_p = 32000$ (top row) and $N_p = 64000$ (bottom row). We compare results for QMC-hybrid runs(left), MC-hybrid runs(middle) and non-hybrid runs(right). The scattering limit N_s is set to zero for hybrid runs and ∞ for non-hybrid runs.