



Addressing nonlinearities in Monte Carlo

Jeremi Dauchet, Jean-Jacques Bézian, Stéphane Blanco, Cyril Caliot, Julien Charon, Christophe Coustet, Mouna El-Hafi, Vincent Eymet, Olivier Farges, Vincent Forest, et al.

► To cite this version:

Jeremi Dauchet, Jean-Jacques Bézian, Stéphane Blanco, Cyril Caliot, Julien Charon, et al.. Addressing nonlinearities in Monte Carlo. Scientific Reports, 2018, 8 (1), art.13302-11 p. 10.1038/s41598-018-31574-4 . hal-01871366

HAL Id: hal-01871366

<https://imt-mines-albi.hal.science/hal-01871366>

Submitted on 10 Sep 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

SCIENTIFIC REPORTS

OPEN

Addressing nonlinearities in Monte Carlo

Jérémi Dauchet¹, Jean-Jacques Beziau², Stéphane Blanco³, Cyril Caliot⁴, Julien Charon^{1,2,4}, Christophe Coustet⁵, Mouna El Haf², Vincent Eymet⁵, Olivier Farges⁶, Vincent Forest⁵, Richard Fournier³, Mathieu Galtier⁷, Jacques Gautrais⁸, Anaïs Khuong⁸, Lionel Pelissier⁹, Benjamin Piaud⁵, Maxime Roger⁷, Guillaume Terrée² & Sebastian Weitz^{1,2}

Received: 19 February 2018

Accepted: 3 August 2018

Published online: 05 September 2018

Monte Carlo is famous for accepting model extensions and model refinements up to infinite dimension. However, this powerful incremental design is based on a premise which has severely limited its application so far: a state-variable can only be recursively defined as a function of underlying state-variables if this function is linear. Here we show that this premise can be alleviated by projecting nonlinearities onto a polynomial basis and increasing the configuration space dimension. Considering phytoplankton growth in light-limited environments, radiative transfer in planetary atmospheres, electromagnetic scattering by particles, and concentrated solar power plant production, we prove the real-world usability of this advance in four test cases which were previously regarded as impracticable using Monte Carlo approaches. We also illustrate an outstanding feature of our method when applied to acute problems with interacting particles: handling rare events is now straightforward. Overall, our extension preserves the features that made the method popular: addressing nonlinearities does not compromise on model refinement or system complexity, and convergence rates remain independent of dimension.

The standard Monte Carlo (MC) method is a technique to predict a physical observable by numerically estimating a statistical expectation over a multi-dimensional configuration space¹. The reason why this method is so popular in all fields of scientific research is its intuitive nature. In general, simulation tools are designed in direct relation to the physical phenomena present in each discipline, and later refinements are gradual and straightforward. Model refinements merely extend sampling to other appropriate dimensions. The method is nonetheless mathematically rigorous: specialists specify observables that are implicitly translated into integral quantities which are estimated using random sampling in each direction of the configuration space. This statistical approach is highly powerful because the algorithm can be designed directly from the description of the system, whether it is deterministic or not, with no reworking or approximation.

Let us illustrate how MC is used in engineering with a typical example: the optimal design of a concentrated solar plant² (see Fig. 1a). The power collected by the central receiver results from all the rays of sunlight that reach it after reflection by heliostats, so it depends on the complex geometry of the heliostats. Moreover, the heliostats change their orientation to follow the sun's position, so they can mask one another at certain times of the day. To estimate by MC the received power at one moment of interest, i.e. for a given geometry of the heliostats: choose an optical path among those that link the sun to the central receiver via a heliostat; check whether this path is shadowed or blocked by another heliostat; and retain a Monte Carlo weight equal to 0 or 1 depending on transmission success. Let X be the random variable denoting transmission success. The collected fraction of the

¹Université Clermont Auvergne, CNRS, SIGMA Clermont, Institut Pascal, F-63000, Clermont-Ferrand, France.

²Université Fédérale de Toulouse Midi-Pyrénées, Mines Albi, UMR CNRS 5302, Centre RAPSODEE, Campus Jarlard, F-81013, Albi CT Cedex, France. ³LAPLACE, Université de Toulouse, CNRS, INPT, UPS, Toulouse, France. ⁴Processes, Materials and Solar Energy Laboratory, PROMES, CNRS, 7 rue du Four Solaire, 66120, Font-Romeu-Odeillo-Via, France. ⁵Méso-Star SAS, LONGAGES, France. ⁶Université de Lorraine, LEMTA, CNRS, UMR 7563, Vandœuvre-lès-Nancy, F-54500, France. ⁷Univ Lyon 1, CNRS, INSA-Lyon, Université Claude Bernard Lyon 1, CETHIL UMR5008, F-69621, Villeurbanne, France. ⁸Centre de Recherches sur la Cognition Animale, Centre de Biologie Intégrative (CBI), Centre National de la Recherche Scientifique (CNRS), Université Paul Sabatier (UPS), F-31062, Toulouse Cedex 9, France. ⁹Education, Formation, Travail, Savoirs (EFTS), Université de Toulouse, ENSFEA, UT2J, Toulouse, France. Correspondence and requests for materials should be addressed to J.D. (email: jeremi.dauchet@sigma-clermont.fr)

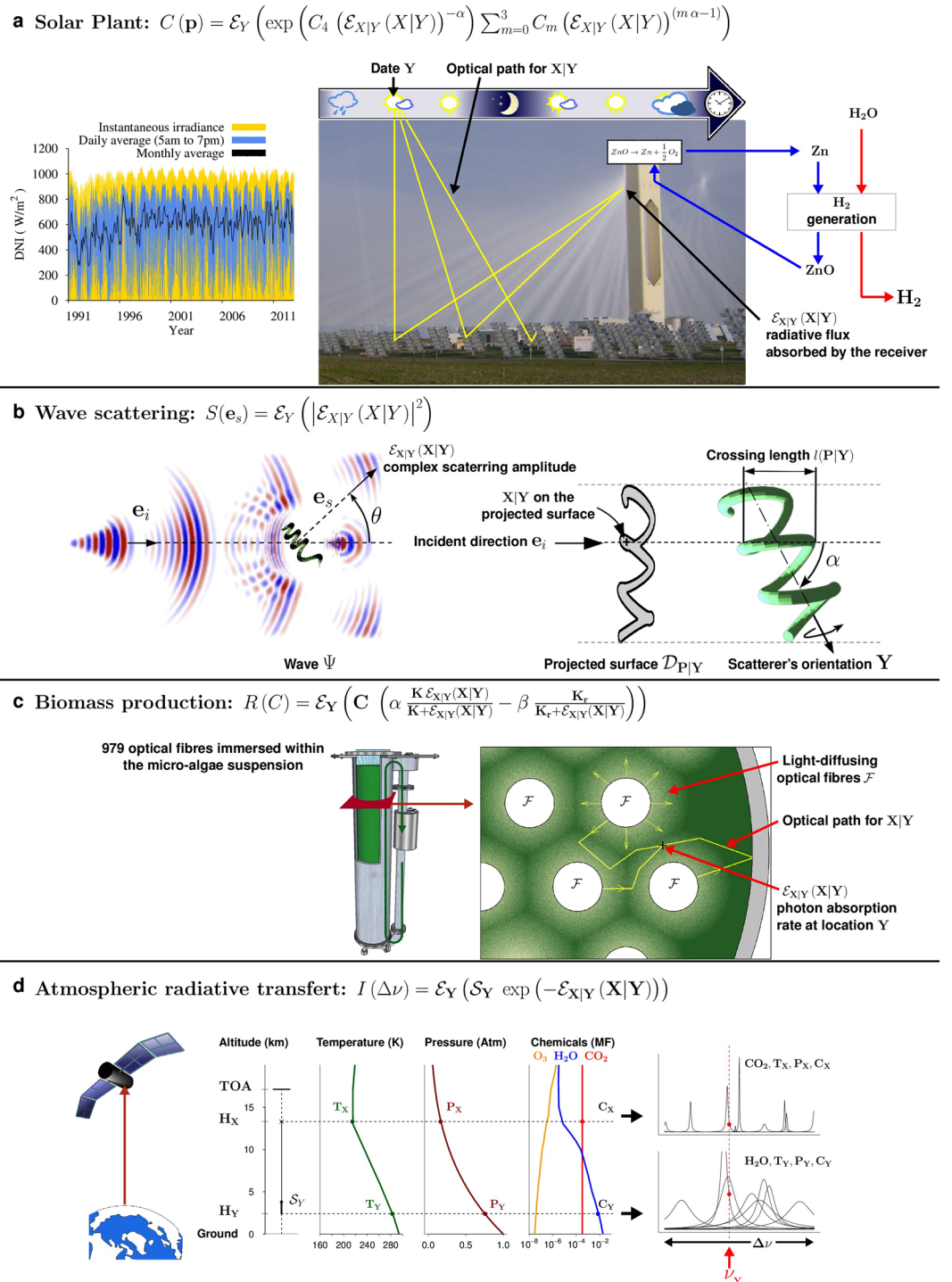


Figure 1. Complex systems with nonlinear outputs: four real-world examples. **(a)** Solar-driven high-temperature thermal reduction of zinc oxide, as the first phase of a two-step water-splitting cycle. Photons emitted from the sun are reflected on heliostats and concentrated at the entrance of the chemical reactor in which ZnO dissociation is carried out. Depending on their transmission success X , the solar power $\mathcal{E}_{X|Y}(X|Y)$ absorbed by the receiver at a random instant Y determines the nonlinear chemical conversion rate of the reaction $\text{ZnO} \rightarrow \text{Zn} + \frac{1}{2}\text{O}_2$. Here we address the estimation of the solar plant's annual conversion rate $C(\mathbf{p})$ at different Earth locations \mathbf{p} , by averaging the instantaneous conversion rates over the statistics of sun position and incident Direct Normal Irradiance (DNI), which fluctuates with time and weather conditions (see also SI1). **(b)** Wave scattering by a complex-shaped and optically soft scatterer (cyanobacterium *Arthrospira*). An incident plane wave with propagation direction \mathbf{e}_i is scattered by the helical cyanobacterium. The bacterium has low relative refractive index and is much larger than the wavelength (optically soft particle). The complex scattering amplitude $\mathcal{E}_{X|Y}(X|Y)$ in the forward directions is the sum of secondary wave contributions $X|Y$ (interference) originating from the scatterer surface. This surface depends on the scatterer orientation Y . Here we address the estimation of $S(\mathbf{e}_s)$, the single-scattering differential cross-section, in direction \mathbf{e}_s for a suspension of particles,

assuming independent scattering, by averaging the squared modulus of $\mathcal{E}_{X|Y}(X|Y)$ over the statistics of orientations Y (see also SI2). (c) Phytoplankton growth in light-limited environments. Phytoplankton is grown in a continuous stirred tank photobioreactor, internally illuminated by optical fibres \mathcal{F} immersed in the culture. The local rate of photon absorption $\mathcal{E}_{X|Y}(X|Y)$ at location Y is the average of the contributions $X|Y$ of every optical path from the fibres to Y through the scattering and absorbing suspension. $\mathcal{E}_{X|Y}(X|Y)$ determines the nonlinear photosynthetic growth rate at location Y . Here we address the Monte Carlo estimation of $R(C)$, the global growth-rate in the whole culture volume, as a function of biomass concentration C , by averaging the local rate over locations in the volume (see also SI3). (d) Atmospheric radiative transfer: top-of-atmosphere (TOA) specific intensity (from earth towards outer space). Photons emitted by the atmosphere at infrared frequencies are due to random emission transitions Y , from a higher to a lower energy state, of mainly CO_2 and H_2O molecules of concentration C_Y at altitude H_Y . The corresponding source \mathcal{S}_Y depends on the thermodynamic state of the atmosphere, mainly temperature T_Y (defining the energy-state population) and pressure P_Y (defining most of the line width, i.e. the uncertainty of the emission frequency ν_Y). This source is then exponentially attenuated by atmospheric absorption, i.e. by all random absorption transitions $X|Y$, from a lower to a higher energy state, occurring at altitude $H_{X|Y}$ between H_Y and the top of the atmosphere (see also SI4). Copyright. “Central solar PS10” (<https://commons.wikimedia.org/wiki/File:Luz.jpg>) by MwAce is released into the public domain. “satellite” (www.pixabay.com/en/satellite-solar-panels-space-297840), “blue-earth” (pixabay.com/en/earth-blue-land-globe-planet-297125), “forecast icons” (www.pixabay.com/en/weather-signs-symbols-forecast-28719) and “time-clock” (pixabay.com/en/clock-time-hour-minute-wall-clock-295201) by Clker-Free-Vector-Images (www.pixabay.com/en/users/Clker-Free-Vector-Images-3736) are licensed under CC0 Creative Commons (www.creativecommons.org/publicdomain/zero/1.0/deed.en).

available sun power is then the expectation $\mathcal{E}_X(X)$ of X , and can be evaluated with no bias as the average of such weights over a large number of sampled paths.

This approach robustly complies with expanded descriptions of the physical observable to be addressed. For instance, the fraction of the available sun power collected on average over the entire lifetime of the solar plant (typically 30 years) can be predicted as the expectation over time of $\mathcal{E}_X(X)$, which varies with time. Denoting $\mathcal{E}_{X|Y}(X|Y)$ the collected fraction at random time Y within the 30 years, the time-averaged fraction is given by $\mathcal{E}_Y(\mathcal{E}_{X|Y}(X|Y)) = \mathcal{E}_{Y,X|Y}(X|Y)$. The basic algorithm above can then be encapsulated within time sampling: first choose a date for Y , then pick a path at that date for $X|Y$. Finally, estimate $\mathcal{E}_{Y,X|Y}(X|Y)$ by computing the average transmission success over all combined pairs (date, path). Meanwhile, sun power fluctuations can be accounted for by estimating the atmospheric transmission at each chosen date. The choice of the statistical viewpoint thus enables us to incorporate into one single statistical question as many elements as necessary: the geometrical complexity of the heliostats³, the daily course of the sun, and seasonal-scale as well as hourly-scale weather fluctuations⁴. Remarkably, the latter question is nearly as simple to address as the estimation of the power collected at one single date: the algorithmic design can map the full conceptual description, yet computational costs are hardly affected. Contrastingly, deterministic approaches would translate into impractical computation times or require simplified and approximate descriptions, so MC has become the only practical solution in many engineering contexts of this type. Having become standard practice, MC has prompted numerous theoretical developments^{5–8}.

Nevertheless, MC has so far not been able to handle *every* question. In fact, it was identified early on that “the extension of Monte Carlo methods to nonlinear processes may be impossible”⁹ and it is a prevalent opinion nowadays that “Monte Carlo methods are not generally effective for nonlinear problems, mainly because expectations are linear in character”¹⁰, so that “a nonlinear problem must usually be linearized in order to use the Monte Carlo technique”¹¹. We are aware of only one attempt so far to bypass this failing: the recent proposal by the applied mathematics community^{1,12–14} to use branching processes¹⁵ to solve Fredholm-type integral equations with polynomial nonlinearity.

Unfortunately, most real-world problems are nonlinear. Indeed, if the question were now to evaluate the final return on investment of the solar plant, namely how much electrical power it would deliver over its lifetime, standard MC would fail, because the instantaneous conversion efficiency from collected solar power to electrical power is not linear. Let us consider, as a toy example, a basic nonlinear case where the electrical power would be proportional to the square of the instantaneous collected solar power $\mathcal{E}_{X|Y}(X|Y)$ at date Y . In Monte-Carlo terms, the question would then be to estimate $\mathcal{E}_Y(\mathcal{E}_{X|Y}(X|Y)^2)$ over the plant’s lifetime. In this case, the optical and temporal expectations can no longer be combined, because it would be wrong to first estimate, as above, the total solar power collected over its lifetime, and then apply the conversion efficiency at the end (basically, $\mathcal{E}_Y(\mathcal{E}_{X|Y}(X|Y)^2) \neq \mathcal{E}_Y(\mathcal{E}_{X|Y}(X|Y))^2$, in the same way as $a^2 + b^2 \neq (a + b)^2$). Instead, we would have to sample dates (say M dates, millions over 30 years), estimate the solar power collected at each date by averaging transmission successes over numerous optical paths (say N paths, millions for each date), apply a nonlinear conversion to the result at that date, and then average over all dates¹⁶. Doing so, MC would now require $M \times N$ samples, and even worse, further levels of complexity (each adding a nonlinearity to the problem) would similarly multiply the computation time. Moreover, the result would be biased due to the finite sampling sizes of the innermost dimensions. In short, MC’s distinctive features are no longer available, and exact lifetime integration appears impossible.

Bearing in mind our earlier theoretical works about MC integral formulations², we have found a way to bypass this obstacle for a large class of nonlinear problems, based on the very statistical nature of MC. In the case of our toy example, we use the fact that:

$$\mathcal{E}_Y(\mathcal{E}_{X|Y}(\mathbf{X}|\mathbf{Y})^2) = \mathcal{E}_{Y,(\mathbf{X}_1, \mathbf{X}_2)|Y}(\mathbf{X}_1 \mathbf{X}_2 | \mathbf{Y}) \quad (1)$$

where \mathbf{X}_1 and \mathbf{X}_2 are two independent variables, identically distributed as \mathbf{X} (see Methods). Translated into a sampling algorithm, the solution is now to sample optical paths in pairs $(\mathbf{X}_1, \mathbf{X}_2)|Y$ (instead of millions) at each sampled date, and then to retain the pair product $\mathbf{X}_1 \mathbf{X}_2 | Y$ of their transmission successes. The optical and temporal statistics can then actually be sampled together, and yield the unbiased result with no combinatorial explosion. This reformulation can be generalised to any nonlinearity of polynomial shape. First, monomials of any degree can indeed be estimated using the same sampling property as that used above for $n = 2$:

$$\mathcal{E}_Y(\mathcal{E}_{X|Y}(\mathbf{X}|\mathbf{Y})^n) = \mathcal{E}_{Y,(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)|Y}(\mathbf{X}_1 \mathbf{X}_2 \dots \mathbf{X}_n | \mathbf{Y}) \quad (2)$$

where \mathbf{X}_i are n independent random variables, identically distributed as \mathbf{X} . For any monomial of degree n , the expectation can then be computed by sampling series of n independent realisations of $\mathbf{X}|\mathbf{Y}$, and averaging the series products. The linear case, solved by standard MC, corresponds to $n = 1$. Secondly, since polynomials are simply linear combinations of monomials, the expectation for any polynomial function of $\mathcal{E}_{X|Y}(\mathbf{X}|\mathbf{Y})$ of degree n can be translated into a Monte Carlo algorithm, first sampling a degree in the polynomial, and then sampling as many independent realisations of $\mathbf{X}|\mathbf{Y}$ as this random degree (see Methods). For a polynomial function of degree n , the corresponding Non-Linear Monte Carlo (NLMC) algorithm is then:

- pick a sample y of Y ,
- choose a monomial degree value $d \leq n$,
- draw d independent samples of $\mathbf{X}|\mathbf{Y} = y$ and retain their product,

repeat this sampling procedure and compute the estimate as the average of the retained products.

Moreover, if polynomial forms of any dimension are now solvable with no approximation, so is the projection of any nonlinear function onto a polynomial basis of any dimension, even of infinite dimension if required (full details of using the Taylor expansion are given in Methods). As a result, any hierarchy of nested statistical processes that combine nonlinearly can now, in theory, be exactly addressed within the Monte Carlo framework. The deep rationale of the proposed algorithm is therefore to transform a nonlinear process into a formally equivalent linear infinite-dimension process, and then use the inherent capability of Monte Carlo to address expectations over domains of infinite dimension.

To the best of our knowledge, this analysis has never before been performed. However, it has major practical consequences for real-world problems, provided the polynomial sampling, which is the price to be paid for tackling nonlinearities exactly, remains tractable. For instance, let us go back to our solar power plant example, and now use the actual expression for the conversion rate and its Taylor expansion: for each date, once a sun position and climate conditions have been fixed, we would have to pick a random number of independent optical paths (instead of one optical path in the linear case), keep the product of their transmission success, and finally calculate the average of many such products. Doing so, it becomes possible to integrate hourly solar input fluctuations over 30 years in the full geometry of a kilometre-wide heliostat field in order to optimise the nonlinear solar-to-electric conversion over the plant lifetime (Fig. 1a). The same line of thought can be used to predict wave scattering by a tiny complex-shaped scatterer¹⁷ such as a helicoidal cyanobacterium (Fig. 1b). The biomass production example (Fig. 1c), where incoming light favours the photosynthetic growth that in turn blocks the incoming light, illustrates how our method also handles nonlinear feedback¹⁸. Finally, with the estimation of Earth's radiative cooling (Fig. 1d), we reproduce quite classic results, yet with a purely statistical approach: by sampling directly the state transitions of greenhouse gases, we avoid costly deterministic computations that the standard linear Monte Carlo approach requires in order to by-pass the nonlinearity of the Beer Extinction Law¹⁹. In each of the four cases, it appears that the additional computations are well affordable using only ordinary computing power (the complete physical descriptions of the four problems, the nonlinearities involved and their translation in NLMC can be found in their respective Extended Data Figures and Supplemental Information, Solar Plant: SI1; Complex-shaped Scatterer: SI2; Biomass production: SI3; Earth radiative cooling: SI4).

For these four real-world simulation examples, we can therefore retain that the variance of the proposed statistical estimate was very much satisfactory. Is that a general feature? Can we feel confident when applying this simulation strategy to any new nonlinear problem? More generally speaking, what do we claim about the status of the present research? Essentially, we only argue that the general proposition of the present paper is immediately available for an ensemble of practical applications. Indeed, these four simulation examples are representative of a quite wide ensemble of physics/engineering practices and the corresponding implementations are now practically used by the corresponding research-communities^{17,19–21}. Moreover, implementation only required an up to date knowledge of Monte Carlo techniques: the probability sets were selected using nothing more than very standard importance-sampling reasoning (see Methods, SI1 and SI3). Outside these experiments, we did not explore in any systematic manner the statistical convergence difficulties that could be specifically associated to the proposition. But although we did not yet encounter it, we can already point out a potential source of variance related to the choice of the fixed point x_0 around which the nonlinear function is Taylor expanded (see Methods).

From a theoretical point of view, in the four cases exposed above, the model is directly enunciated in statistical terms, defining two random variables \mathbf{X} and \mathbf{Y} from the start. More broadly, standard MC practice can also start from a deterministic description (see Methods), most commonly from a linear partial differential equation (PDE). The formal equivalence between the solution of a linear PDE and the expectation of a random variable has long been established²². Indeed, PDE-to-MC translations are essential to nanoscale mechanics (Quantum Monte Carlo²³) or nuclear sciences. NLMC allows such translations for nonlinear PDEs.

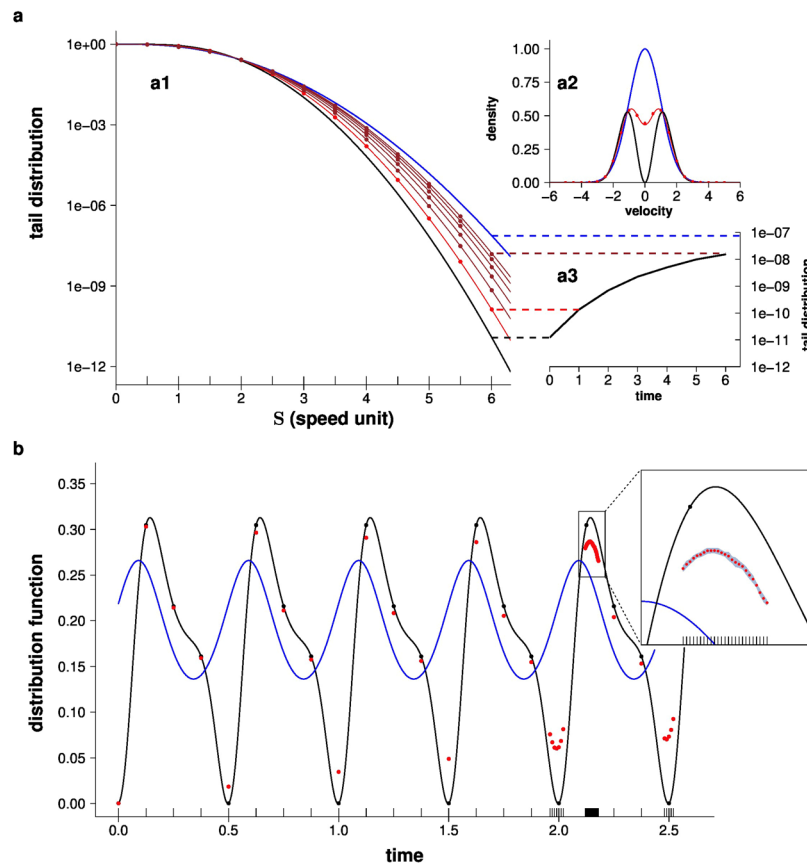


Figure 2. Nonlinear Monte Carlo for gas kinetics. **(a)** Relaxation of speed distribution to equilibrium ((a1) tail distribution of particle speeds (fraction of particles faster than S), (a2) probability density function of velocity, (a3) variation in the fraction of particles faster than 6 speed units). In a homogeneous gas, collisions between particles redistribute velocities so that the speed distribution tends to equilibrium (Maxwellian distribution). Starting from a distribution far from equilibrium (black curves), we compute the relaxation to a Maxwellian distribution (blue curves) by estimating the tail distribution at different times (e.g. red curves correspond to the system state at 1 unit time). The continuous lines correspond to the analytical solutions and each point corresponds to an independent NLMC estimation. The fraction of particles which are faster than 6 speed units (a3) illustrates how NLMC well accounts for the 1000-fold increase in the rarest high-speed particles, with no space or time discretisation. Remarkably, rare events are estimated with the same relative uncertainty as frequent events (10^4 samples for each estimate; confidence intervals of all estimates are contained within the thickness of the point). **(b)** Dampening of breathing mode. A dilute gas confined by an outside harmonic potential displays a *breathing mode*. We estimate the density at probe position (1.75, 0, 0) and velocity (0.35, 0, 0) (in adimensional units, see SI5) at different times (adim. unit). Starting from a distribution complying with the local equilibrium, the density displays a perpetual oscillation at twice the trap frequency (blue curve, analytical solution), independent of the collision rate (or, equivalently, of the elastic cross section). Starting from a distribution which is far from the local equilibrium (the same initial distribution as in (a)), the density still pulsates in the absence of collisions (null cross section; black curve: analytical solution; black points: probe estimates). When collisions are introduced (by raising the cross section), the velocity redistribution induced by the collisions dampens the oscillation (red points: probe estimates; no analytical solution). The inset illustrates that probes can be bunched together to zoom in on a period of special interest (e.g. estimating the peak values at each cycle). Each point is estimated independently using 10^7 samples; confidence intervals of all estimates are contained within the thickness of one point in the main figure, and confidence intervals are represented by the grey background area in the inset).

As an illustration of the ground-breaking nature of our study, we address a prominent example of a nonlinear PDE in statistical physics, the Boltzmann equation, which governs the spatiotemporal density of interacting particles in a dilute gas (full details in SI5). The corresponding physics is easy to visualise: a particle simply follows its ballistic flight until it collides with another particle. The collisions are considered as instantaneous and only modify the two particle velocities. The equation for the variation in particle density in phase-space (position, velocity) is nonlinear because the collision rate depends on the density itself. In order to project this nonlinearity onto the proper polynomial basis of infinite dimension, this PDE is first translated into its Fredholm integral counterpart (a step reminiscent of the aforementioned Dimov proposition¹). This Fredholm integral expresses the density in phase space at some location for some velocity at some time, as if putting a probe into space-time.

It is estimated by Monte Carlo, tracking the dynamics backwards in time up to the initial condition (or boundary conditions). Importantly, such a probe estimation does not require the exhaustive resolution of the whole field at previous times: as in standard backwards MC algorithms for solving linear transport (e.g. simulating an image by tracking photon-paths backward, from receiver to source^{24–26}) the information about previous states of the field is reconstructed along each path only where and when it is required²⁷. Here, the contrast with linear MC is that nonlinearity due to collisions translates into branching paths.

This extension deals very efficiently with extremely rare events because it preserves an essential feature of MC: by avoiding time/space/velocity discretisation^{28–30}, very low densities can be estimated with no bias, and the only source of uncertainty is the finite number of sampled events (i.e. the confidence interval around the estimated density). As a test, we consider a case for which analytical solutions have been published: Krook's early analysis of the distribution of speeds in extremely out-of-equilibrium conditions^{31,32}. Krook's analysis was outstanding because it provided an analytical solution to a problem which looked impossible to solve numerically: events with the greatest consequences, namely the particles with the highest energies (i.e. high-speed particles, of tremendous importance in nuclear chemistry) lie far out in the tail of speed distribution and have a very low probability of occurrence (rare events). Using our NLMC design, the fractions of particles which have a kinetic energy higher than 10^6 times the average value, and which correspond to a fraction as low as 10^{-11} of the total, can now be quantified as accurately as desired, and perfectly fit the analytical solution (Fig. 2a).

Having been validated in Krook's case, this extension opens the way to solving systems for which no analytical solutions are available. As an example, we now consider a fully spatialised system in which the particles are confined by an outside harmonic potential, leading to a so-called *breathing mode* of the gas density. Such a solution to the Boltzmann equation was identified early on by Boltzmann himself³³, but has recently been revisited and generalised in the context of a shortcut to adiabaticity techniques for classic gases³⁴. Exact solutions are available only under the constraint that the gas is at local equilibrium, in which case the density displays a permanent oscillation. Here again, these analytical solutions are exactly recovered. Moreover, NLMC enables us to go beyond this constraint and to explore the gas behaviour when the local equilibrium constraint is alleviated: starting from a state far from local equilibrium, it is now possible to estimate how fast the velocity redistribution induced by collisions actually dampens the oscillation (Fig. 2b).

Conclusions

From now on, the Monte Carlo Method is no longer restricted to linear problems. The five examples exposed above were worked out by teams comprising specialists of the Monte Carlo method and specialists of the physical problem under consideration. Through their complete description, we offer readers all the details to implement their own applications. As a guideline, the first step is to formulate the physical observable under its expectation form, including the nonlinearities and integrating all levels of complexity. The second step is to reformulate this expectation as a formulation compliant with the standard Monte Carlo Method, according to the type of nonlinearity. For polynomial nonlinearities, use i.i.d. series products. For other differentiable forms, use a Taylor expansion around an upper bound of the innermost random variable in order to regain a polynomial form. Using this MC-compliant formulation, every advanced MC technique can then be applied: parallel implementation, complex geometry, null collisions, zero variance, control variate, importance sampling, sensitivity analysis, and so on. As illustrated by the variety of our seminal examples, this guideline covers a large set of nonlinear academic and real-world problems.

Methods

Basics of Monte Carlo Methods. Let us estimate $E = 1 + 4$ by repeatedly tossing a (fair) coin. The tossing process is described by a random variable $R \in \{0, 1\}$ which takes the value 0 for heads (probability $P_R(0) = \frac{1}{2}$) and 1 for tails (probability $P_R(1) = \frac{1}{2}$).

Now, to estimate any process (e.g. a process output: $E = 1 + 4$), we can assign arbitrary weights $w(R)$ to values $\{0, 1\}$ in order to write E as an expectation of the weighted process, following:

$$E = 1 + 4 = P_R(0)w(0) + P_R(1)w(1) = \mathcal{E}_R(w(R)) \quad (3)$$

with $w(0) = \frac{1}{P_R(0)} = 2$ and $w(1) = \frac{4}{P_R(1)} = 8$ and where \mathcal{E}_R denotes the expectation with respect to R . Using the results $r_1 \dots r_N$ of N successive tosses (independent realisations of R), we can then estimate $E = \mathcal{E}_R(w(R))$ from the weighted average of the toss results $\frac{1}{N} \sum_{i=1}^N w(r_i)$ since $E = 5$ is indeed the average of Monte Carlo weights that take the values 2 and 8 with equal probabilities.

Such an approach is at the base of Monte Carlo techniques: define the weights according to the problem to be solved, sample the process repeatedly, and take the average. Depending on the physical description of the value to be estimated, this approach still holds for an infinite number of terms and can also be extended to integral formulation using continuous random variables:

$$E_Y(w(Y)) = \int_{\mathcal{D}_Y} dy P_Y(y) w(y) \quad (4)$$

which can be estimated by $\frac{1}{N} \sum_{i=1}^N w(y_i)$, where the y_i are N realisations of the random variable Y with probability density function p_Y and domain of definition \mathcal{D}_Y .

Basics of Nonlinear Monte Carlo Methods. In order to estimate

$$E = \mathcal{E}_Y(\mathcal{E}_{X|Y}(\mathbf{X}|\mathbf{Y})^2) = \int_{\mathcal{D}_Y} dy p_Y(y) \left(\int_{\mathcal{D}_{X|Y}} dx p_{X|Y}(x|y) x \right)^2 \quad (5)$$

we introduce two independent variables \mathbf{X}_1 and \mathbf{X}_2 , identically distributed as \mathbf{X} (still conditioned by the same \mathbf{Y}):

$$\begin{aligned} E &= \mathcal{E}_Y(\mathcal{E}_{X_1|Y}(\mathbf{X}_1|\mathbf{Y}) \mathcal{E}_{X_2|Y}(\mathbf{X}_2|\mathbf{Y})) \\ &= \int_{\mathcal{D}_Y} dy p_Y(y) \left(\int_{\mathcal{D}_{X_1|Y}} dx_1 p_{X_1|Y}(x_1|y) x_1 \right) \left(\int_{\mathcal{D}_{X_2|Y}} dx_2 p_{X_2|Y}(x_2|y) x_2 \right) \end{aligned} \quad (6)$$

Since \mathbf{X}_1 and \mathbf{X}_2 are independent, and conditionally independent given \mathbf{Y} :

$$\begin{aligned} E &= \int_{\mathcal{D}_Y} dy p_Y(y) \left(\iint_{\mathcal{D}_{X_1|Y}^2} dx_1 p_{X_1|Y}(x_1|y) dx_2 p_{X_2|Y}(x_2|y) x_1 x_2 \right) \\ &= \mathcal{E}_Y(\mathcal{E}_{(X_1, X_2)|Y}(\mathbf{X}_1 \mathbf{X}_2|\mathbf{Y})) \end{aligned} \quad (7)$$

Hence

$$\begin{aligned} E &= \iiint_{\mathcal{D}_Y \times \mathcal{D}_{X_1|Y}^2} dy p_Y(y) dx_1 p_{X_1|Y}(x_1|y) dx_2 p_{X_2|Y}(x_2|y) x_1 x_2 \\ &= \mathcal{E}_{Y, (X_1, X_2)|Y}(\mathbf{X}_1 \mathbf{X}_2|\mathbf{Y}) \end{aligned} \quad (8)$$

The same demonstration can be made to establish that:

$$\mathcal{E}_Y(\mathcal{E}_{X|Y}(\mathbf{X}|\mathbf{Y})^n) = \mathcal{E}_{Y, (X_1, X_2, \dots, X_n)|Y}(\mathbf{X}_1 \mathbf{X}_2 \dots \mathbf{X}_n|\mathbf{Y}) \quad (9)$$

Let us now assume that the weights associated with the random variable \mathbf{Y} are described by a nonlinear function $\mathbf{f}(\mathbf{Z}_Y)$ of the conditional expectation $\mathbf{Z}_Y = \mathcal{E}_{X|Y}(\mathbf{X}|\mathbf{Y})$. The problem then becomes to compute:

$$E = \mathcal{E}_Y(\mathbf{f}(\mathbf{Z}_Y)) = \mathcal{E}_Y(\mathbf{f}(\mathcal{E}_{X|Y}(\mathbf{X}|\mathbf{Y}))) \quad (10)$$

Such a nonlinearity can be treated with no approximation using a projection on an infinite basis. In all the examples presented in this article, we have used a Taylor polynomials basis, which means that $\mathbf{f}(x)$ is expanded around x_0

$$\mathbf{f}(x) = \sum_{n=0}^{+\infty} \frac{\partial^n \mathbf{f}(x_0)}{n!} (x - x_0)^n \quad (11)$$

We note that both x_0 and \mathbf{f} can be conditioned by \mathbf{Y} . Now, following the same line as explained above for the Basics of Monte Carlo Methods, we regard the sum in the expansion of \mathbf{f} as an expectation, writing:

$$\mathbf{f}(x) = \mathcal{E}_N \left(\frac{\partial^N \mathbf{f}(x_0)}{P_N(N)N!} (x - x_0)^N \right) \quad (12)$$

where the random variable N (of probability law P_N) is the degree of one monomial in the Taylor polynomial. This step only requires us to define one infinite set of probabilities (instead of two in Eq. 3), with $\sum_{n=0}^{+\infty} P_N(n) = 1$.

Equation 10 can then be rewritten as:

$$E = \mathcal{E}_Y(\mathbf{f}(\mathcal{E}_{X|Y}(\mathbf{X}|\mathbf{Y}))) = \mathcal{E}_{Y, N} \left(\frac{\partial^N \mathbf{f}(x_0)}{P_N(N)N!} (\mathcal{E}_{X|Y}(\mathbf{X}|\mathbf{Y}) - x_0)^N \right) \quad (13)$$

Defining independent and identically distributed random variables \mathbf{X}_q , with the same distribution as \mathbf{X} , the innermost term rewrites

$$E = \mathcal{E}_{Y, N} \left(\frac{\partial^N \mathbf{f}(x_0)}{P_N(N)N!} \prod_{q=1}^N (\mathcal{E}_{X_q|Y}(\mathbf{X}_q|\mathbf{Y}) - x_0) \right) \quad (14)$$

or, equivalently:

$$E = \mathcal{E}_{Y, N} \left(\frac{\partial^N \mathbf{f}(x_0)}{P_N(N)N!} \prod_{q=1}^N \mathcal{E}_{X_q|Y}(\mathbf{X}_q|\mathbf{Y} - x_0) \right) \quad (15)$$

Since the variables $\mathbf{X}_q|\mathbf{Y}$ are independent in the innermost term, we have:

$$\prod_{q=1}^N \mathcal{E}_{\mathbf{X}_q|Y}(\mathbf{X}_q|Y - x_0) = \mathcal{E}_{(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)|Y} \left(\prod_{q=1}^N (\mathbf{X}_q|Y - x_0) \right) \quad (16)$$

so that:

$$E = \mathcal{E}_{Y,N} \left(\frac{\partial^N f(x_0)}{P_N(N)N!} \mathcal{E}_{(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)|Y} \left(\prod_{q=1}^N (\mathbf{X}_q|Y - x_0) \right) \right) \quad (17)$$

and we finally have:

$$E = \mathcal{E}_{Y,N,(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)|Y} \left(\frac{\partial^N f(x_0)}{P_N(N)N!} \prod_{q=1}^N (\mathbf{X}_q|Y - x_0) \right) \quad (18)$$

which can be read as:

$$E = \mathcal{E}_{Y,N,(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)|Y}(w(Y, N, \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)) \quad (19)$$

with

$$w(Y, N, \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N) = \frac{\partial^N f(x_0)}{P_N(N)N!} \prod_{q=1}^N (\mathbf{X}_q|Y - x_0) \quad (20)$$

With the notation above, $\prod_{q=1}^0 (\mathbf{X}_q|Y - x_0) = 1$.

The translation into a Monte Carlo algorithm then follows:

- sample a realisation y of Y (and set x_0 and f accordingly if they depend on y)
- sample a realisation n of N
- sample n independent realisations $x_{q=1, \dots, n}$ of \mathbf{X} conditioned by y
- keep

$$\hat{w} = w(y, n, x_1, \dots, x_n) = \frac{\partial^n f(x_0)}{P_N(n)n!} \prod_{q=1}^n (x_q - x_0)$$

and estimate E as the average of many realisations \hat{w} .

Implementation example. Let us illustrate the choice of the discrete distribution P on N with an implementation example. We take Y uniformly distributed over $[0, 1]$, $X|Y$ uniformly distributed over $[0, Y]$ and $f(x) = 1/(1+x)$ (f corresponds to the photobioreactor real-world example in Fig. 1c, with $C = 1$, $\alpha = 0$, $\beta = -1$, $K_r = 1$). Equation 10 becomes

$$E = \mathcal{E}_Y \left(\frac{1}{1 + \mathcal{E}_{X|Y}(X|Y)} \right) \quad (21)$$

Its analytical solution is $E = 2 \ln(3/2)$.

Injecting the n -th derivative $\partial^n f(x_0) = \frac{n! (-1)^n}{(1+x_0)^{n+1}}$ into Equation 20 leads to

$$w(Y, N, X_1, X_2, \dots, X_N) = \frac{(-1)^N}{P_N(N) (1+x_0)^{N+1}} \prod_{q=1}^N (X_q|Y - x_0)$$

that can be reformulated as

$$w(Y, N, X_1, X_2, \dots, X_N) = \frac{1}{P_N(N) (1+x_0)^{N+1}} \prod_{q=1}^N \frac{X_q|Y - x_0}{x_0} \quad (22)$$

Using standard importance-sampling reasoning, we choose the set of probabilities that cancels the term $\frac{x_0^N}{(1+x_0)^{N+1}}$ in the estimator:

$$w(Y, N, X_1, X_2, \dots, X_N) = \prod_{q=1}^N \frac{X_q|Y - x_0}{x_0} \quad (23)$$

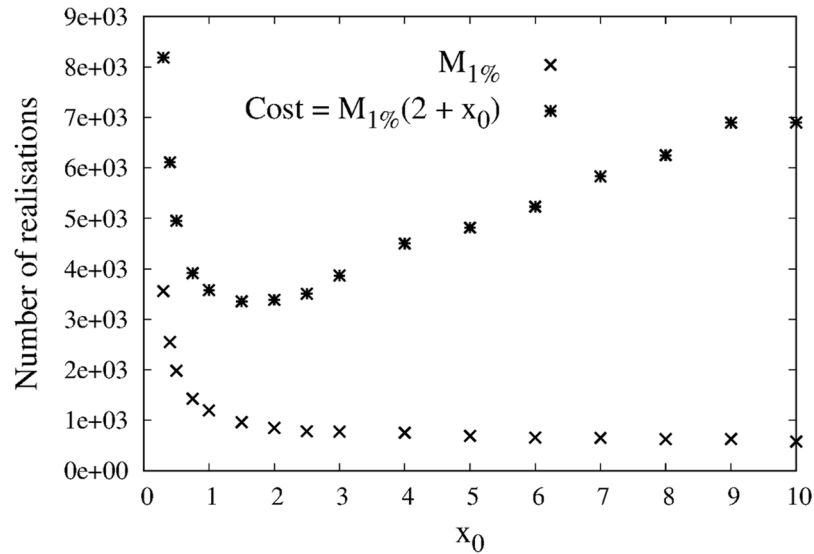


Figure 3. Choice of the discrete distribution P on N : a tradeoff between estimation and computational cost. Estimation of equation 21 with the estimator in equation 23, as a function of the choice of the fixed point x_0 around which the nonlinear function is Taylor expanded. In this example, x_0 also defines the discrete distribution P_N (see equation 24). Crosses indicate the number of realisations $M_{1\%}$ that are required to achieve an estimation with 1% standard deviation. Stars indicate the corresponding computational cost as defined in equation 25.

with

$$P_N(N) = \frac{x_0^N}{(1 + x_0)^{N+1}} \quad (24)$$

The NLMC algorithm is

- sample a realisation y of Y
- sample a realisation n of N according to the discrete distribution in Equation 24
- sample n independent realisations $x_{q=1,\dots,n}$ of X uniformly distributed over $[0, y]$
- keep

$$\hat{w} = \prod_{q=1}^n \frac{x_q - x_0}{x_0}$$

and estimate E as the average of M realisations \hat{w} .

We define the computational cost C of this algorithm in terms of the total number of random generations that are required to achieve 1% standard deviation on the estimation. Each realisation of the algorithm includes 1 random generation y of Y , 1 random generation n of N , and n random generations of X . And it takes $M_{1\%}$ realisations of the algorithm to achieve a standard deviation of 1%. Overall,

$$C = M_{1\%}(2 + \mathcal{E}(N)) = M_{1\%}(2 + x_0) \quad (25)$$

since $\mathcal{E}(N) = x_0$ with the discrete distribution in Equation 24. Figure 3 shows the values of $M_{1\%}$ and C recorded with simulations, as a function of x_0 . The choice of x_0 alone controls both the statistical convergence (*i.e.* $M_{1\%}$) and the computational cost (through the discrete distribution P on N). We observe a trade-off between estimation and computational cost. For low values of x_0 , only few realisations of X are needed since the discrete distribution on N is rapidly decreasing with n , but a large number of realisation of the algorithm are required for the estimation (*i.e.* $\mathcal{E}(N)$ is small but $M_{1\%}$ is large). At the opposite, for larger values of x_0 , the estimator converges rapidly, but the average number of X random generations per Monte Carlo realisations is increased (*i.e.* $M_{1\%}$ is small but $\mathcal{E}(N)$ is large). In between, we observe an optimal choice of x_0 .

Comparison with the naive plug-in estimator and convergence issues. In the previous implementation example solving equation 21, a naive plug-in estimator could be constructed as¹⁶:

$$E = \mathcal{E}_Y \left(\frac{1}{1 + \frac{1}{K} \sum_{q=1}^K X_q | Y} \right)$$

leading to the following Monte Carlo algorithm:

- sample a realisation y of Y
- sample K independent realisations $x_{q=1,\dots,K}$ of X uniformly distributed over $[0, y]$
- keep

$$\hat{w} = \frac{1}{1 + \frac{1}{K} \sum_{q=1}^K x_q}$$

and estimate E as the average of M realisations \hat{w} . With this algorithm, K must ensure that the bias of the estimator can be neglected. For that purpose, we choose the value $K_{1\%}$ that always gives an estimation of $\mathcal{E}_{X|Y}$ with 1% standard deviation. Therefore, each realisation of the algorithm includes 1 random generation for Y and $K_{1\%}$ generations for X , and it takes $M_{1\%}$ realisations of the algorithm to achieve a standard deviation of 1% on the estimation of E . The computational cost of the naive algorithm is therefore $M_{1\%}(1 + K_{1\%})$. In the present example, $K_{1\%} = 3333$ and we observed that $M_{1\%} = 140$ thanks to numerical simulations: the cost is $C = 466760$. Compared to the results in Fig. 3, even for this very simple example where Y and $X|Y$ have little variance, the computational cost of the naive plug-in algorithm is 100 times higher than that of the NLMC estimator. Nevertheless, this conclusion only stands for a reasonable choice of x_0 (and therefore of P_N). Indeed, the computational cost with the NLMC estimator seems to rise up to infinity when x_0 approaches 0 (see Equation 23 and Fig. 3): even the naive plug-in algorithm would then be a better choice. Although we did not further analyse this observation with theoretical means, we can at least retain that choosing x_0 is likely to be an essential step of the present approach as far as computational costs are concerned.

References

1. Dimov, I. T. & McKee, S. Monte Carlo Methods for Applied Scientists (World Scientific Publishing, 2008).
2. Delatorre, J. *et al.* Monte Carlo advances and concentrated solar applications. *Sol. Energy* **103**, 653–681 (2014).
3. Siala, F. M. F. & Elayeb, M. E. Mathematical formulation of a graphical method for a no-blocking heliostat field layout. *Renew. energy* **23**, 77–92 (2001).
4. Farges, O. *et al.* Life-time integration using Monte Carlo Methods when optimizing the design of concentrated solar power plants. *Sol. Energy* **113**, 57–62 (2015).
5. Assaraf, R. & Caffarel, M. Zero-variance principle for Monte Carlo algorithms. *Phys. Rev. Lett.* **83**, 4682 (1999).
6. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**, 1087–1092 (1953).
7. Hammersley, J. M. & Handscomb, D. C. *Monte Carlo Methods*. (Springer, Netherlands, 1964).
8. Roger, M., Blanco, S., El Hafi, M. & Fournier, R. Monte Carlo Estimates of Domain-Deformation Sensitivities. *Phys. Rev. Lett.* **95**, 180601 (2005).
9. Curtiss, J. H. 'Monte Carlo' Methods for the Iteration of Linear Operators. *J. Math. Phys.* **32**, 209–232 (1953).
10. Kalos, M. H. & Whitlock, P. A. *Monte Carlo Methods*. second ed., (Wiley-VCH, Weinheim, 2008).
11. Chatterjee, K., Roadcap, J. R. & Singh, S. A new Green's function Monte Carlo algorithm for the solution of the two-dimensional nonlinear Poisson–Boltzmann equation: Application to the modeling of the communication breakdown problem in space vehicles during re-entry. *J. Comput. Phys.* **276**, 479–485 (2014).
12. Vajargah, B. F. & Moradi, M. Monte Carlo algorithms for solving Fredholm integral equations and Fredholm differential integral equations. *Appl. Math. Sci.* **1**, 463–470 (2007).
13. Rasulov, A., Raimova, G. & Mascagni, M. Monte Carlo solution of Cauchy problem for a nonlinear parabolic equation. *Math. Comput. Simulation* **80**, 1118–1123 (2008).
14. Gobet, E. Monte-Carlo Methods and Stochastic Processes: From Linear to Non-linear (CRC Press, 2016).
15. Skorokhod, A. V. Branching diffusion processes. *Theory Probab. Appl.* **9**, 445–449 (1964).
16. Hong, L. J. & Juneja, S. Estimating the Mean of a Non-linear Function of Conditional Expectation. Proceedings of the 2009 Winter Simulation Conference, Austin, Texas, 1223–1236 (2009).
17. Charon, J. *et al.* Monte Carlo implementation of Schiff's approximation for estimating radiative properties of homogeneous, simple-shaped and optically soft particles: Application to photosynthetic micro-organisms. *J. Quant. Spectrosc. Radiat. Transf.* **172**, 3–23 (2016).
18. Cornet, J. F. Calculation of optimal design and ideal productivities of volumetrically lightened photobioreactors using the constructal approach. *Chem. Eng. Sci.* **65**, 985–998 (2010).
19. Galtier, M. *et al.* Radiative transfer and spectroscopic databases: A line-sampling Monte Carlo approach. *J. Quant. Spectrosc. Radiat. Transf.* **172**, 83–97 (2016).
20. Dauchet, J. *et al.* Calculation of the radiative properties of photosynthetic microorganisms. *J. Quant. Spectrosc. Radiat. Transfer*. (2015).
21. Dauchet, J., Cornet, J.-F., Gros, F., Roudet, M. & Dussap, C.-G. Chapter One – Photobioreactor Modeling and Radiative Transfer Analysis for Engineering Purposes. *Adv. Chem. Eng.* **48**, 1–106 (2016).
22. Kac, M. On some connections between probability theory and differential and integral equations. Proc. Second Berkeley Symp. Math. Statistics Probab. 189 (1951).
23. Corney, J. F. & Drummond, P. D. Gaussian quantum Monte Carlo methods for fermions and bosons. *Phys. Rev. Lett.* **93**, 2–5 (2004).
24. Pharr, M. & Humphreys, G. Physically Based Rendering: from theory to implementation (Elsevier, 2010).
25. Case, K. M. Transfer problems and the reciprocity principle. *Rev. Mod. Phys.* **29**, 651 (1957).
26. Collins, D. G., Blättner, W. G., Wells, M. B. & Horak, H. G. Backward Monte Carlo calculations of the polarization characteristics of the radiation emerging from spherical-shell atmospheres. *Appl. Opt.* **11**, 2684–2696 (1972).
27. Galtier, M. *et al.* Integral formulation of null-collision Monte Carlo algorithms. *J. Quant. Spectrosc. Radiat. Transf.* **125**, 57–68 (2013).
28. Wagner, W. Stochastic particle methods and approximation of the Boltzmann equation. *Math. Comput. Simul.* **38**, 211–216 (1995).
29. Rjasanow, S. A Stochastic Weighted Particle Method for the Boltzmann Equation. *J. Comput. Phys.* **124**, 243–253 (1996).
30. Rjasanow, S. & Wagner, W. Simulation of rare events by the stochastic weighted particle method for the Boltzmann equation. *Math. Comput. Model.* **33**, 907–926 (2001).
31. Krook, M. & Wu, T. T. Formation of Maxwellian Tails. *Phys. Rev. Lett.* **36**, 1107–1109 (1976).
32. Krook, M. & Wu, T. T. Exact solutions of the Boltzmann equation. *Phys. Fluids* **20**, 1589–1595 (1977).
33. Boltzmann, L. In *Wissenschaftliche Abhandlungen*, edited by Hasenorl, F. Vol. II, p. 83 (J.A. Barth, Leipzig, 1909).
34. Guéry-Odelin, D., Muga, J. G., Ruiz-Montero, M. J. & Trizac, E. Nonequilibrium Solutions of the Boltzmann Equation under the Action of an External Force. *Phys. Rev. Lett.* **112**, 180602 (2014).

Acknowledgements

The authors express their deep gratitude to Igor Roffiac for fruitful discussions on the Monte Carlo method. This work was sponsored by the French National Centre for Scientific Research (CNRS) through the PEPS-JCJC OPTISOL_Mu program, by the French Agence Nationale de la Recherche (ANR) under grant ANR-16-CE01-0010 (project High-Tune), by the Region Occitanie under grant CLE-2016-EDStar and by the French government research-program “Investissements d’avenir” through the LABEXs ANR-10-LABX-16-01 IMobS3 and ANR-10-LABX-22-01 SOLSTICE and the ATS program ALGUE of IDEX ANR-11-IDEX-02 UNITI.

Author Contributions

All authors contributed extensively to the theoretical developments presented in this paper. Each author contributed to the practical applications according to his or her scientific expertise: J.D., M.E.H., V.E., R.F. and M.G. in Atmospheric sciences, J.D., S.B., C.Ca., M.E.H., V.E., O.F., R.F., M.G. and M.R. in Radiative Transfer, J.D., S.B., R.F., J.G., A.K. and S.W. in Complex Systems in Biology, J.D., J.J.B., S.B., C.Ca., M.E.H., V.E., O.F. and R.F. in Solar Energy, J.D., S.B., J.C., M.E.H. and R.F. in Electromagnetism and quantum mechanics, J.D., S.B., M.E.H., R.F., J.G., B.P. and G.T. in Statistical Physics. C.Co., V.E., V.F. and B.P. (www.meso-star.com) performed the numerical implementations. J.D., S.B., R.F. and J.G. wrote the paper.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-31574-4>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018