Radiative transfer and spectroscopic databases: A line-sampling Monte Carlo approach

Mathieu Galtier, Stephane Blanco, Jérémi Dauchet, Mouna El-Hafi, Vincent Eymet, Richard Fournier, Maxime Roger, Christophe Spiesser, Guillaume Terrée

Radiative transfer and spectroscopic databases: a line-sampling Monte Carlo approach

M. Galtier[a,b,*], S. Blanco[d,e], J. Dauchet[f,g], M. El Hafi[b], V. Eymet[c], R. Fournier[d,e], M. Roger[a], C. Spiesser[b], G. Terrée[b]

[a]*Université de Lyon, CNRS, CETHIL, UMR5008, F-69621 Villeurbanne, France*
[b]*Université de Toulouse, Mines-Albi, UMR-CNRS 5302, Centre RAPSODEE, Campus Jarlard, F-81013 Albi CT cedex 09, France*
[c]*Méso-Star SAS, 8 rue des pêchers, 31410 Longages, France*
[d]*Université de Toulouse, UPS, LAPLACE (Laboratoire Plasma et Conversion d'Énergie), 118 route de Narbonne, F-31062 Toulouse cedex 9, France*
[e]*CNRS, LAPLACE, F-31062 Toulouse, France*
[f]*Clermont Université, ENSCCF, Institut Pascal, BP 10448, F-63000 Clermont-Ferrand, France*
[g]*CNRS, UMR 6602, IP, F-63171 Aubière, France*

Abstract

Dealing with molecular-state transitions for radiative transfer purposes involves two successive steps that both reach the complexity level at which physicists start thinking about statistical approaches: 1/ constructing line-shaped absorption spectra as the result of very numerous state-transitions, 2/ integrating over optical-path domains. For the first time, we show here how these steps can be addressed simultaneously using the null-collision concept. This opens the door to the design of Monte Carlo codes directly estimating radiative transfer observables from spectroscopic databases. The intermediate step of producing accurate high-resolution absorption spectra is no longer required. A Monte Carlo algorithm is proposed and applied to six one-dimensional test cases. It allows the computation of spectrally integrated intensities (over 25cm^{-1} bands or the full IR range) in a few seconds, regardless of the retained database and line model. But free parameters need to be selected and they impact the convergence. A first possible selection is provided in full detail. We observe that this selection is highly satisfactory for quite distinct atmospheric and combustion configurations, but a more systematic exploration is still in progress.

Keywords: Radiative transfer, Monte Carlo method, Null-collision, Line sampling, Statistical approach, Spectroscopic databases

Nomenclature

We retain here the following standard statistical formalism: random variables are written in upper-case (*e.g. X*), free and bound variables in lower-case (*e.g. x*) and samples of random variables are subscripted (*e.g.* x_m).

A_i Bernoulli random variable

B_ν Blackbody intensity

---

*Corresponding author. Tel.: +33 4 72 43 71 68
Email address: `mathieu.galtier@insa-lyon.fr` (M. Galtier)

d	Distance
E	Statistical expectation
$h_{a,v,j}$	Contribution of transition j to the absorption coefficient
$\hat{h}_{v,j}$	Contribution of transition j to the extinction coefficient
i	Index of collision
i_b	Narrowband index
I_v	Monochromatic specific intensity
j	Index of molecular transition
$k_{a,v}$	Absorption coefficient
$k_{n,v}$	Null-collision coefficient
$k_{s,v}$	Scattering coefficient
$\hat{k}_v$	Extinction coefficient resulting from null-collision introduction
$\tilde{k}_v$	Approximate absorption coefficient
m	Index of independent realization
M	Number of independent realizations
N_b	Number of narrowbands
N_t	Number of molecular transitions
P_a	Probability of absorption
$P_{a,J}(j)$	Probability of absorption for a given transition j
$P_b(i_b)$	Probability of narrowband i_b
$P_J(j)$	Probability of transition j
P_n	Probability of null collision
$p_X(x)$	Probability density function for absorption/emission location
$\hat{p}_X(x)$	Probability density function for absorption/emission location considering null-collision events
$p_N(v)$	Probability density function for wavenumber
$p_s(\omega\|\omega')$	Single scattering phase function
$Q_J(j)$	Second component of the proposed transition probability expression
r	Random number sampled uniformly over the unit interval
S_k	Moments of order k
t	Time
T	Temperature

$t_{1\%}$	Computation time require to get a 1% standard deviation
$T_U(u)$	First component of the proposed transition probability expression
u	Index of spectral bands
VOV	Variance of variance
w	Monte Carlo weight
x	Location (scalar)
x_0	Observation location
$\mathbf{x}$	Location (vector)
y	Absorption/emission location at the end of a path involving null-collisions
α	Multiplicative factor
γ_Q	Half-width of an arbitrary Lorentz profile
δ	Delta-Dirac distribution
δ_π	Partial derivative with respect to parameter π
$\delta\nu$	Spectral interval
ζ	Extinction threshold
ν	Wavenumber
π	Generic parameter
τ	Transmissivity
ω	Direction (vector)

1. Introduction

Recent advances in Monte Carlo methods [1, 2] indicate that null-collision algorithms can be used to simulate radiative transfer in semi-transparent media in such a way that grids are no longer required. Even for highly non-homogeneous configurations the volume does not need to be discretized and this is achieved while still preserving all the statistical properties of standard Monte Carlo algorithms: convergence toward the exact solution is rigorous (no bias) and for a finite number M of samples, the estimate is associated with a faithful statistical uncertainty ($\frac{1}{\sqrt{M}}$ times the standard deviation of the Monte Carlo weights). The initial idea of null-collisions came from the plasma physics and neutron transport communities. It consists in adding a third type of collisional event to the absorption and scattering events. These new collisions change nothing with regard to the propagation of radiation. They are pure-forward scattering events: after the collision, the photon continues its trajectory as if no collision had occurred (the collision is "null" or "virtual"). The amount of such collisions can therefore be tuned to any value without modifying the resulting transfer. This tuning can then be such that the total density of the collisionners is uniform: more null-collisions are added where true extinction is weak, fewer where it is strong, so that the total extinction coefficient is uniform (as in a homogeneous semi-transparent medium). The consequence is that collision locations can be sampled within a simple

exponential-distribution, according to the Beer extinction law. It is thus only at the sampled location that the question is raised of evaluating the respective amounts of absorbers, true-scatterers and null-collisionners, in order to sample one type of events out of the three possible ones (according to their respective probabilities). The photon-transport Monte Carlo algorithm is therefore decoupled from the spatial description of radiative properties. It is not required to "know" the true radiative properties when transporting photons until their next collision location.

In preceding null-collision algorithms, the radiative properties were indeed known and the only benefit was a simplification of the procedure for accessing the data (getting rid of the volume-discretization requirement). Here we assume that the monochromatic absorption coefficient $k_{a,\nu}$ is unknown and that the only available data are those of the molecular spectroscopic databases [3, 4, 5, 6]. We still have $k_{a,\nu} = \sum_{j=1}^{N_t} h_{a,\nu,j}$, where N_t is the number of molecular transitions and $h_{a,\nu,j}$ is the known contribution of the j-th transition to the absorption coefficient at wavenumber ν, but this sum is not precomputed. Null-collisions allow us to initiate the Monte Carlo algorithm without computing $k_{a,\nu}$ and we then observe that the algorithm can be continued without ever making this computation: $\sum_{j=1}^{N_t} h_{a,\nu,j}$ is diluted into the photon-transport algorithm via successive samplings of the transition index.

Why this is possible is quite simple. Let us take only two transitions: $k_{a,\nu} = h_{a,\nu,1} + h_{a,\nu,2}$ and write $k_{a,\nu} = P_1 \frac{h_{a,\nu,1}}{P_1} + P_2 \frac{h_{a,\nu,2}}{P_2}$ with P_1 and P_2 in $]0, 1[$ and $P_1 + P_2 = 1$. This justifies a Monte Carlo algorithm sampling the index j of one of the two transitions, $j = 1$ with probability P_1, $j = 2$ with probability P_2, and evaluating $k_{a,\nu}$ as the average of M weights $w(j) = \frac{h_{a,\nu,j}}{P_j}$ (a Bernoulli trial algorithm for computing sums). With this procedure, the absorption coefficient is viewed as the expectation of a random process. But in null-collision algorithms $k_{a,\nu}$ appears only within the absorption probability, in a pure linear manner. So the process of sampling optical paths can be combined with the sampling of transitions: once a path has been sampled, at the stage of summing the contributions of transition 1 and transition 2 to evaluate $k_{a,\nu}$, only one transition is sampled and $w(j)$ is used in place of $k_{a,\nu}$ (the expectation is replaced by the value of only one sampled event). Instead of two successive algorithms, one for the evaluation of $k_{a,\nu}$ and the other for radiative transfer, a single higher dimension Monte Carlo algorithm is built that combines the two integration spaces (illustrated in the following paragraphs).

This idea is particularly meaningful when dealing with very large numbers of transitions and when the question of pre-computing the absorption coefficient is an issue. For the engineering applications that have motivated the most recent developments of Monte Carlo algorithms based on line-by-line approaches [7, 8, 9, 10], the issue usually does not arise: for each new version of molecular spectroscopic databases, look-up tables are constructed that allow, via interpolation procedures, most of the gaseous mixtures and thermodynamic states encountered in industrial devices to be covered. We are thinking more in terms of study domains where spectroscopic requirements for either accuracy or compositional diversity are more critical. This is the case for the terrestrial atmosphere (LW applications for global radiative budget modeling [11] or SW radiative transfer when dealing with remote-sensing [12, 13]) and astrophysical studies (when studying atmospheres of exoplanets [14], brown dwarfs [15, 16], or stellar atmospheres [17, 18]). In these contexts, constructing look-up tables for each new configuration or each new spectroscopic assumption is a very

demanding intermediate step and the approach introduced in the present article allows us to envisage the complete disappearance of this step.

The practical value of this new approach can be illustrated with the simple example of evaluating the spectral average transmissivity of a non-homogeneous and non-scattering gaseous column. Using the null-collision algorithm of [1] (assuming that $k_{a,\nu}$ is precomputed) would lead to the following w-sampling algorithm:

- Step 1 : Uniform sampling of a wavenumber ν within the considered spectral band.

- Step 2 : Sampling of a distance d traveled until the next collision as if the medium was homogeneous of extinction coefficient $\hat{k}_{\nu}$ (this distance is sampled according to $d = \frac{1}{\hat{k}_{\nu}} ln(r)$ where r is sampled uniformly on the unit interval).

- Step 3 : If the traveled distance leads the photon to exit the column, the algorithm stops with $w = 1$ (transmission).

- Step 4 : If the traveled distance leads the photon to a collision location within the column, the available absorption data are accessed to evaluate $k_{a,\nu}$ at the collision location and the absorption probability P_a is computed as $P_a = \frac{k_{a,\nu}}{\hat{k}_{\nu}}$.

- Step 5 : A Bernoulli trial of probability P_a is performed to retain an absorption event or a null-collision event (absorption with probability P_a, null-collision with probability $1 - P_a$).

- Step 6 : If absorption, the algorithm stops with $w = 0$ (no transmission).

- Step 7 : If null-collision, the algorithm loops at Step 2.

This sampling algorithm is repeated M times to get M values of w, indexed w_m, each equal either to 0 or 1, and the spectral average transmissivity is evaluated as $\tau \approx \frac{1}{M} \sum_{m=1}^{M} w_m$. The only free parameter is the value of $\hat{k}_{\nu}$, only affecting the computation speed, and not the convergence speed, provided that $\hat{k}_{\nu} > k_{a,\nu}$ [1]. Alternatively, the present approach leads to the following slightly modified algorithm (that will be fully established in Sec. 2):

- Step 1 : Uniform sampling of a wavenumber ν within the considered spectral band.

- Step 2 : Sampling of a distance d traveled until next collision as if the medium was homogeneous of extinction coefficient $\hat{k}_{\nu}$.

- Step 3 : If the traveled distance leads the photon to exit the column, the algorithm stops with $w = 1$ (transmission).

---

[1]The contribution of null-collisions to the total-extinction coefficient is $\hat{k}_{\nu} - k_{a,\nu}$ and $\hat{k}_{\nu} < k_{a,\nu}$ would imply a negative value of the density of null-collisionners. See the discussion in [1]

- Step 4 : If the traveled distance leads the photon to a collision location within the column, a transition j is sampled among the N_t available transitions, thanks to a Bernoulli trial of probabilities $P_1, P_2...P_{N_t}$.

- Step 5 : A Bernoulli trial of probability $P_{a,j} = \frac{h_{a,\nu,j}}{\hat{k}_\nu P_j}$ is performed to retain an absorption event or a null-collision event (absorption with probability $P_{a,j}$, null-collision with probability $1 - P_{a,j}$).

- Step 6 : If absorption, the algorithm stops with $w = 0$ (no transmission).

- Step 7 : If null-collision, the algorithm loops at Step 2.

The only difference here is that instead of using the known value of $k_{a,\nu}$ to compute the absorption probability as $\frac{k_{a,\nu}}{\hat{k}_\nu}$, a transition j is first sampled and only the contribution of this sampled transition $h_{a,\nu,j}$ is required to compute $P_{a,j}$. The recursive algorithm is then continued identically. Figure 1 displays the results of a first academic implementation test that was made in order to check the feasibility of such Monte Carlo simulations when dealing with the typical distances and inhomogeneities encountered in atmospheric applications (strong variations of composition and pressure along the vertical with line widths more than ten times smaller at the top than at the bottom of the atmosphere). Each simulation point in Figure 1 required between 0.1s and 12s on a standard laptop and the results are in perfect adequacy with solutions obtained independently using a standard deterministic line-by-line simulation (using a 50-levels discretization of pressure and composition along the vertical).
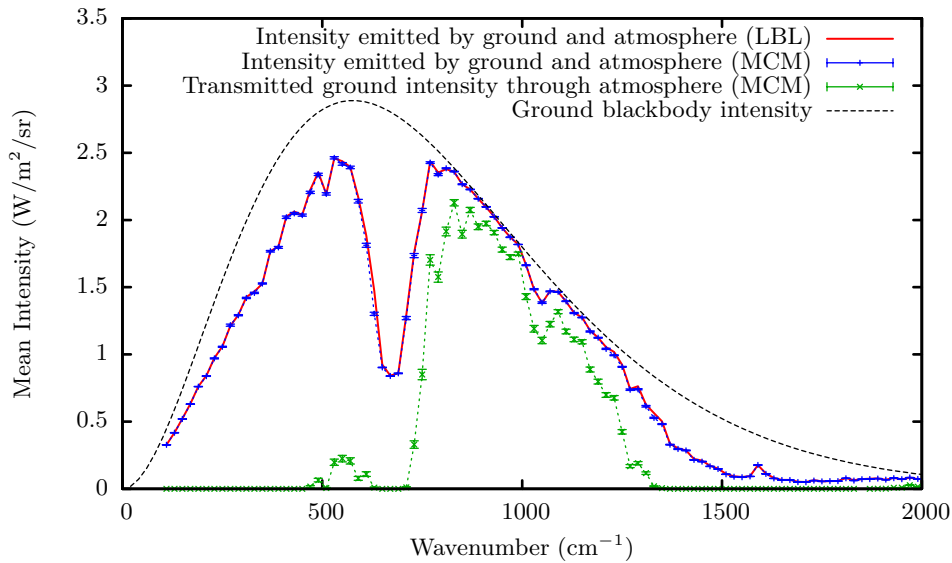

Figure 1: Upward vertical intensities, averaged over $20cm^{-1}$ bands, at altitude $17km$ for the mid-latitude summer atmosphere [19]. Simulation times correspond to computations performed using an Intel Core i7 processor (2.8GHz).

This test was sufficient to convince us that the approach could be practical and that it would at least be useful to researchers wishing to bypass the task of constructing look-up tables when briefly testing a new spectroscopic assumption or testing the influence of new lines in an updated version of a molecular spectroscopic database (before

upgrading the look-up tables accordingly). However, we are only at the preliminary stage of fully validating the approach, which involves providing:

- a complete theoretical proof establishing that these algorithms (combining the sampling of the optical-path with the sampling of transitions) do indeed evaluate the addressed radiative quantities in an unbiased manner;

- a validation against published results on configurations including all the required spectral complexity (with significant impacts of line-profile assumptions as well as spectroscopic-database alternatives).

These are the objectives of the following sections. Sec. 2 provides all the details of the required statistics, the resulting Monte Carlo algorithm allowing the computation of spectrally-integrated intensities is then described in Sec. 3 and is tested in Sec. 4 against the simulation results of [20]. Finally, in Sec. 5 some considerations about the meaning of such statistical pictures outside Monte Carlo applications are discussed and a conclusion (Sec. 6) makes several comments on the proposed approach from an operational viewpoint.

2. Statistical formulation

This section gradually establishes a statistical formulation combining random transitions with random collision-locations. This leads to the writing of the specific-intensity as the expectation of a random weight that the Monte Carlo algorithm will sample M times in order to build a specific-intensity estimate. We start by recalling the null-collision concept (Sec. 2.1) as well as standard statistical pictures associated with the estimation of specific-intensity in non-scattering infinite media (Sec. 2.2). We then modify this picture, once by introducing null-collisions (Sec. 2.3), and twice by splitting the absorption coefficient into transition contributions (Sec. 2.4).

2.1. The null-collision concept

As shown in [1], introducing null-collisions consists in reformulating the radiative transfer equation

$$\frac{\partial I_\nu}{\partial t} + \omega.\nabla I_\nu = -(k_{a,\nu} + k_{s,\nu})I_\nu + k_{a,\nu}B_\nu + k_{s,\nu}\int_{4\pi} I'_\nu p_s(\omega|\omega')d\omega' \tag{1}$$

into

$$\frac{\partial I_\nu}{\partial t} + \omega.\nabla I_\nu = -\hat{k}_\nu I_\nu + k_{a,\nu}B_\nu + k_{s,\nu}\int_{4\pi} I'_\nu p_s(\omega|\omega')d\omega' + k_{n,\nu}\int_{4\pi} I'_\nu \delta(\omega - \omega')d\omega' \tag{2}$$

where

- $I_\nu \equiv I_\nu(\mathbf{x}, \omega, t)$ is the specific intensity at location $\mathbf{x}$, propagation direction ω, time t and wavenumber ν while I'_ν refers to the same quantity but for propagation direction ω': $I'_\nu \equiv I_\nu(\mathbf{x}, \omega', t)$,

- $B_\nu \equiv B_\nu(\mathbf{x}, t)$ is the blackbody intensity at the local temperature,

- $k_{a,\nu} \equiv k_{a,\nu}(\mathbf{x}, t)$ and $k_{s,\nu} \equiv k_{s,\nu}(\mathbf{x}, t)$ are respectively the absorption and scattering coefficients,

- $p_s(\omega|\omega') \equiv p_{s,\nu}(\mathbf{x}, \omega|\omega', t)$ is the single scattering phase function,

- $\delta(\omega - \omega')$ is the delta Dirac distribution centered along the direction ω.

In Eq. 2, a positive null-collision coefficient $k_{n,\nu} \equiv k_{n,\nu}(\mathbf{x}, t)$ is added to the absorption and scattering coefficients, which means that the initial extinction coefficient $k_{a,\nu} + k_{s,\nu}$ is increased to any arbitrary value

$$\hat{k}_\nu \equiv \hat{k}_\nu(\mathbf{x}, t) = k_{a,\nu} + k_{s,\nu} + k_{n,\nu} \tag{3}$$

Correspondingly, as announced in the introduction, the physical pictures are modified. Instead of having only three types of statistical events (absorption, emission and scattering), a new type is introduced: null-collision events. These events have no effect on radiative transfer since they strictly correspond to pure forward-scattering events (see the delta-Dirac phase function in the fourth term of the right-hand member of Eq. 2) that only compensate the artificial increase of the extinction coefficient. The first immediate benefit of this reformulation is that the new extinction term $-\hat{k}_\nu I_\nu$ no longer depends on the absorption coefficient, but only on $\hat{k}_\nu$. In the configurations studied in [1], the true extinction along a given path was difficult to handle because of large heterogeneities of the temperature and concentration fields, leading to highly varying values of the absorption coefficient. It was much simpler to use a higher but uniform $\hat{k}_\nu$ field instead . Here we are considering configurations for which true extinction is difficult to handle for another reason: independently of any heterogeneity, the gaseous-absorption coefficient is the sum of numerous transition contributions.

2.2. Integral formulation of specific intensity

The discussion is now restricted to an academic configuration. We consider the monochromatic stationary specific intensity at a given location, in a given direction, within an infinite non-scattering medium. Because scattering is neglected, the problem becomes one-dimensional. The addressed intensity resulting from all emissions along the x axis (at all locations between $-\infty$ and $x = x_0$) is noted hereafter as $I_\nu(x_0)$. Eq. 1 then becomes

$$\frac{dI_\nu}{dx} = -k_{a,\nu} I_\nu + k_{a,\nu} B_\nu \tag{4}$$

and its solution at x_0 can be written

$$I_\nu(x_0) = \int_{-\infty}^{x_0} dx \, k_{a,\nu}(x) B_\nu(x) \exp\left(-\int_x^{x_0} k_{a,\nu}(x') dx'\right) \tag{5}$$

which is commonly interpreted, in statistical terms, as the expectation of the blackbody-intensity random function $B_\nu(X)$ where X is a random variable distributed on $]-\infty, x_0]$ according to the probability density $p_X(x) \equiv p_{X,\nu}(x) = k_{a,\nu}(x) \exp\left(-\int_x^{x_0} k_{a,\nu}(x') dx'\right)$:

$$I_\nu(x_0) = E(B_\nu(X)) = \int_{-\infty}^{x_0} p_X(x) \, dx \, B_\nu(x) \tag{6}$$

The corresponding physical picture is that of photons followed from x_0, in the backward direction, until they are absorbed at locations x. Using the reciprocity principle, one then translates these absorption events into emissions

at local temperatures. This leads to $I_\nu(x_0)$ viewed as an average of blackbody intensities $B_\nu(x)$ at backward-sampled locations x (see Fig. 2).
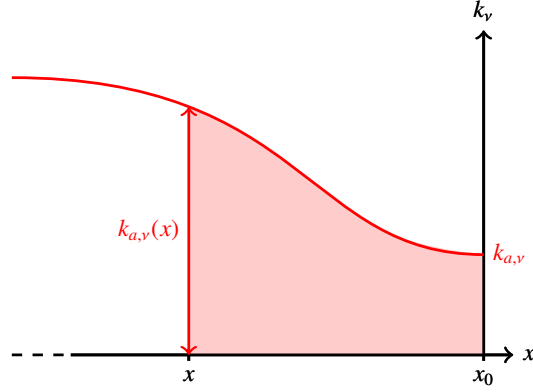

Figure 2: Photons are followed from x_0 in the backward direction until they are absorbed at locations x according to the probability density function $p_X(x) = k_{a,\nu}(x) \exp\left(-\int_x^{x_0} k_{a,\nu}(x')dx'\right)$. Through the reciprocity principle, these absorptions are translated into emissions at local temperatures. In this case, the integral $\int_x^{x_0} k_{a,\nu}(x')dx'$ of the absorption coefficient (colored area) must be available in order to sample absorption/emission locations.

2.3. Introducing null-collisions

This picture can be easily modified to take into consideration the forward-scattering effects of null-collisions. In our simplified one-dimensional configuration, Eq. 2 becomes

$$\frac{dI_\nu}{dx} = -\hat{k}_\nu I_\nu + k_{a,\nu}B_\nu + k_{n,\nu}I_\nu \tag{7}$$

where $\hat{k}_\nu = k_{a,\nu} + k_{n,\nu}$ and its solution at x_0 can be written recursively, as introduced in [1], using

$$I_\nu(x_i) = \int_{-\infty}^{x_i} dx_{i+1} \exp\left(-\int_{x_{i+1}}^{x_i} \hat{k}_\nu(x'_{i+1})dx'_{i+1}\right) \left[k_{a,\nu}(x_{i+1})B_\nu(x_{i+1}) + k_{n,\nu}(x_{i+1})I_\nu(x_{i+1})\right] \tag{8}$$

Statistically speaking, this recursion can be rewritten as

$$I_\nu(x_i) = \int_{-\infty}^{x_i} \hat{p}_{X_{i+1}}(x_{i+1})dx_{i+1} \left[P_a(x_{i+1})B_\nu(x_{i+1}) + P_n(x_{i+1})I_\nu(x_{i+1})\right] \tag{9}$$

where

- the successive collision locations $X_i \equiv X_i(X_{i-1})$ are random variables depending on X_{i-1} and distributed on $]-\infty, X_{i-1}]$ according to the probability density $\hat{p}_{X_i}(x_i) \equiv \hat{p}_{X_i,\nu}(x_i|X_{i-1}) = \hat{k}_\nu(x_i) \exp\left(-\int_{x_i}^{X_{i-1}} \hat{k}_\nu(x'_i)dx'_i\right)$,

- $X_0 = x_0$ is the observation location,

- $P_a(X_i) \equiv P_{a,\nu}(X_i) = k_{a,\nu}(X_i)/\hat{k}_\nu(X_i)$ is the probability of absorption and $P_n(X_i) \equiv P_{n,\nu}(X_i) = 1 - P_a(X_i)$ the probability of null-collision.

9

With these statistical pictures, $I_\nu(x_0)$ is interpreted as the expectation of the random function $B_\nu(Y)$:

$$I_\nu(x_0) = E(B_\nu(Y)) \tag{10}$$

where

- Y is the absorption/emission location at the end of a path involving successive null-collisions. Formally speaking: $Y = \sum_{i=1}^{+\infty} X_i A_i \prod_{q=1}^{i-1}(1 - A_q)$,

- $A_i \equiv A_i(X_i)$ is a Bernoulli random variable, defined as a function of the random variable X_i, taking the value 1 with probability $P_a(X_i)$ and 0 with probability $P_n(X_i)$.

The corresponding physical picture is again that of photons followed from $X_0 = x_0$, in the backward direction, until they are absorbed at locations Y. But the path is now a (pure-forward) multiple-scattering one. At the first collision location X_1, there is a probability $P_a(X_1)$ that absorption occurs. In this case, the path ends at $Y = X_1$. Otherwise, a null-collision occurs, with probability $P_n(X_1)$, and the path is continued (with the same direction) until the next collision at X_2, *etc.* (see Fig. 3).
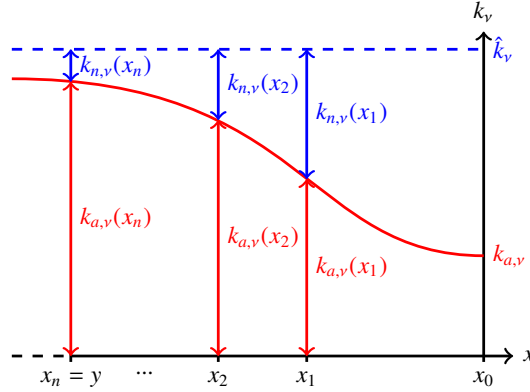

Figure 3: Photons are followed from x_0 in the backward direction until they are absorbed at locations y along a multiple-scattering path (with pure-forward scattering events). At locations x_i there is a probability $P_a(x_i) = k_{a,\nu}(x_i)/\hat{k}_\nu(x_i)$ that absorption occurs. In this case, the path ends at $y = x_i$. Otherwise, a null-collision occurs, with probability $P_n(x_i) = 1 - P_a(x_i)$, and the path is continued until an absorption occurs. In this case, the absorption coefficient $k_{a,\nu}(x_1)$, $k_{a,\nu}(x_2)$, ..., $k_{a,\nu}(x_n)$ at the sampled locations x_1, x_2, ..., x_n must be available in order to sample absorption/emission locations.

2.4. Removal of the absorption coefficient

Considering our objective, the important point at this stage (using the null-collision concept) is that the absorption coefficient does not appear any more in the definition of the successive random locations X_i: they are now defined using the total-extinction $\hat{k}_\nu$ field. $k_{a,\nu}$ only appears in the expression of the probabilities of absorption $P_a(X_i)$ versus

null-collision $P_n(X_i)$ defining the Bernoulli random variable A_i. The multiple-collisions structure of the paths and the nonlinear nature of Beer extinction, i.e. the two main sources of difficulties when analyzing radiative transfer, can be thought through with only $\hat{k}_\nu$ in mind, without any consideration for molecular transitions.

The last step consists in transforming A_i so that $k_{a,\nu}$ vanishes and only transitions are used in the definition of the random variable Y. The absorption coefficient $k_{a,\nu}$ is the sum of the contributions $h_{a,\nu,j}$ of each transition j among N_t transitions[2]. These contributions are $h_{a,\nu,j} = \eta \sigma_{j,\nu}$ with η the molecular density, $\sigma_{j,\nu}$ the cross-section of transition j and $k_{a,\nu} = \sum_{j=1}^{N_t} h_{a,\nu,j}$. We similarly split the null-collision coefficient into contributions of each transition: $k_{n,\nu} = \sum_{j=1}^{N_t} h_{n,\nu,j}$, as if we were now thinking in terms of "null-transitions". The $h_{n,\nu,j}$ values can be chosen arbitrarily, provided that they are positive and that their sum recovers $k_{n,\nu}$. Equation 7 then becomes

$$\frac{dI_\nu}{dx} = -\hat{k}_\nu I_\nu + \sum_{j=1}^{N_t} h_{a,\nu,j} B_\nu + \sum_{j=1}^{N_t} h_{n,\nu,j} I_\nu \tag{11}$$

and its solution at x_0 can be written recursively using

$$I_\nu(x_i) = \int_{-\infty}^{x_i} dx_{i+1} \exp\left(-\int_{x_{i+1}}^{x_i} \hat{k}_\nu(x'_{i+1})dx'_{i+1}\right) \sum_{j_{i+1}=1}^{N_t} \left[h_{a,\nu,j_{i+1}}(x_{i+1})B_\nu(x_{i+1}) + h_{n,\nu,j_{i+1}}(x_{i+1})I_\nu(x_{i+1}) \right] \tag{12}$$

With these new definitions and pictures (illustrated in Fig. 4), when a collision occurs at x_i, it is due to a transition
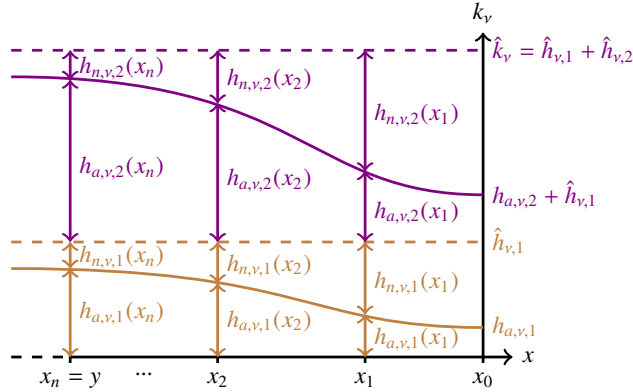

Figure 4: Photons are followed from x_0 in the backward direction until they are absorbed at locations y along a multiple-scattering path (with pure-forward scattering events). At locations x_i, where collision occurs, there is a probability $P_{J_i}(j_i) = \hat{h}_{\nu,j_i}(x_i)/\hat{k}_\nu(x_i)$ that the collision is due to the transition j_i among the N_t possible transitions (the present figure is an illustration for $N_t = 2$). For the transition j_i (which is responsible for collision), there is a probability $P_{a,J_i}(x_i) = h_{a,\nu,j_i}(x_i)/\hat{h}_{\nu,j_i}(x_i)$ that absorption occurs. In this case, the path ends and $y = x_i$. Otherwise, a "null-transition" occurs, with probability $P_{n,J_i}(x_i) = 1 - P_{a,J_i}(x_i)$, and the path is continued until an absorption occurs. In this case, the absorption coefficient completely vanishes: only the transition contributions $h_{a,\nu,j_1}(x_1), h_{a,\nu,j_2}(x_2), ..., h_{a,\nu,j_n}(x_n)$ at the sampled locations $x_1, x_2, ..., x_n$ must be available.

j_i that can be interpreted as a realization of a random variable J_i distributed on $\{1, 2, ..., N_t\}$ with probability $P_{J_i}(j_i) \equiv$

---

[2]For didactic purposes, we have restricted our analysis to a mono-molecular gas. The extension to several species is straightforward, involving only an additional sum over molecules.

$P_{J_i,v}(j_i, x_i) = \hat{h}_{v,j_i}(x_i)/\hat{k}_v(x_i)$, where $\hat{h}_{v,j_i} = h_{a,v,j_i} + h_{n,v,j_i}$ and therefore $\hat{k}_v(x_i) = \sum_{j_i=1}^{N_t} \hat{h}_{v,j_i}(x_i)$. This collision can be either a real one (the photon is absorbed, due to transition j_i) or a null-collision. In statistical terms, Eq. 12 can be rewritten as

$$I_v(x_i) = \int_{-\infty}^{x_i} \hat{p}_{X_{i+1}}(x_{i+1})dx_{i+1} \sum_{j_{i+1}=1}^{N_t} P_{J_{i+1}}(j_{i+1})\left[P_{a,J_{i+1}}(x_{i+1})B_v(x_{i+1}) + P_{n,J_{i+1}}(x_{i+1})I_v(x_{i+1})\right] \tag{13}$$

where $P_{a,J_i}(X_i) \equiv P_{a,v,J_i}(X_i) = h_{a,v,j_i}(X_i)/\hat{h}_{v,j_i}(X_i)$ and $P_{n,J_i}(X_i) \equiv P_{n,v,J_i}(X_i) = h_{n,v,j_i}(X_i)/\hat{h}_{v,j_i}(X_i)$ are respectively the probabilities of absorption and of null-collision for a given transition j_i. With these statistical pictures, $I_v(x_0)$ is interpreted as the expectation of the random function $B_v(Y)$:

$$I_v(x_0) = E(B_v(Y)) \tag{14}$$

where

- Y is again the absorption/emission location, at the end of a path involving successive null-collisions, the only difference being that the sampling of this path is more complex as it now involves the sampling of transitions. Formally speaking we still have: $Y = \sum_{i=1}^{+\infty} X_i A_i \prod_{q=1}^{i-1}(1 - A_q)$,

- The Bernoulli random variable $A_i \equiv A_{i,v}(X_i, J_i)$ now depends on X_i and J_i and takes the value 1 with probability $P_{a,J_i}(X_i)$, 0 with probability $P_{n,J_i}(X_i)$.

3. Monte Carlo

The statistics of the preceding section define a Monte Carlo algorithm that we describe in Sec. 3.1: this is very close to the algorithm that we used for illustration in Sec. 1, with small modifications made to deal with thermal emission instead of only addressing column transmissivities. This simple algorithm is first extended to address spectrally-integrated intensities in Sec. 3.2 and is further refined in Sec. 3.3 (using energy partitioning). Finally, Sec. 3.4 is devoted to the definition of free parameters required by the resulting algorithm that could play an important role in convergence rates.

3.1. Algorithm evaluating a monochromatic intensity

In Eq. 14, $I_v(x_0)$ is the expectation of a random variable $W_v = B_v(Y)$. So translating Eq. 14 into a Monte Carlo algorithm only requires sampling a large number M of independent realizations $\{w_{v,1}, w_{v,2}...w_{v,M}\}$ of W_v, and to use $\tilde{I}_v(x_0) = \frac{1}{M} \sum_{m=1}^{M} w_{v,m}$ as an estimator of $I_v(x_0)$. The most straightforward translation of Eq. 14 leads to the W_v-sampling algorithm introduced in Fig. 5. This first algorithm is easy to follow and easy to implement. It is quite similar to the multiple-scattering algorithms most commonly reported in the literature; only a molecular-transition index is additionally sampled at each (pure forward) scattering event.
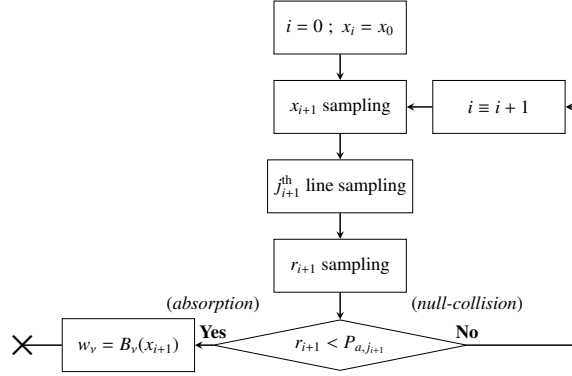
12

Figure 5: The W_ν-sampling algorithm directly translated from Eq. 14. Starting from location x_0 where the intensity is estimated, a location x_1 of collision and a line j_1 are successively sampled. Then by uniformly sampling a random number r_1 between 0 and 1, the type of collision is defined. If $r_1 < P_{a,j_1}(x_1)$ the collision is an absorption, w_ν is computed and the loop stops; otherwise null-collision occurs and the algorithm loops at the step of collision-location sampling. This algorithm is only valid if $\hat{h}_{\nu,j_i} \geq h_{a,\nu,j_i}$ for all locations, lines and wavenumbers. This requirement is needed to make sure that $P_{a,j_i} \in [0,1]$.

3.2. Algorithm evaluating a spectrally integrated intensity

We have slightly modified the previous algorithm by considering the spectrally integrated intensity $I(x_0) = \int_{\nu_{min}}^{\nu_{max}} I_\nu(x_0)d\nu$ instead of the monochromatic intensity. This requires that an additional sampling procedure is used to first select a wavenumber ν on $[\nu_{min}, \nu_{max}]$. $I(x_0)$ is then viewed as the expectation of a new random variable W according to

$$I(x_0) = \int_{\nu_{min}}^{\nu_{max}} E(B_\nu(Y))d\nu = \int_{\nu_{min}}^{\nu_{max}} p_N(\nu)d\nu \left[\frac{1}{p_N(\nu)}E(B_\nu(Y))\right] = E(W) \tag{15}$$

with

$$W = \frac{1}{p_N(N)}B_N(Y) \tag{16}$$

where N is a wavenumber random variable with probability density function $p_N(\nu)$ defined on $[\nu_{min}, \nu_{max}]$ (see Fig. 6). In the preceding section, the values of $\hat{k}$ (or $P_{J_i}(j_i)$) and h_{n,ν,j_i} were arbitrary choices; here $p_N(\nu)$ is also arbitrary: any strictly positive probability density on $[\nu_{min}, \nu_{max}]$ can be chosen.

3.3. Further refinements

We then make two refinements: an energy-partitioning approach is used in order to increase the convergence rate (it switches to Russian roulette at an extinction threshold ζ as proposed in [2]) and an alternative absorption-probability $P_{a,J_i} = \frac{h_{a,\nu,j_i}}{h_{a,\nu,j_i}+|\hat{h}_{\nu,j_i}-h_{a,\nu,j_i}|}$ is introduced to make sure that $P_{a,J_i} \in [0,1]$ whatever the value of $\hat{h}_{\nu,j_i}$ (the solution remains exact because W is modified accordingly, as advised in [1]). The null-collision probability is still defined as $P_{n,J_i} = 1 - P_{a,J_i}$. The resulting Monte-Carlo algorithm is described in Appendix A.
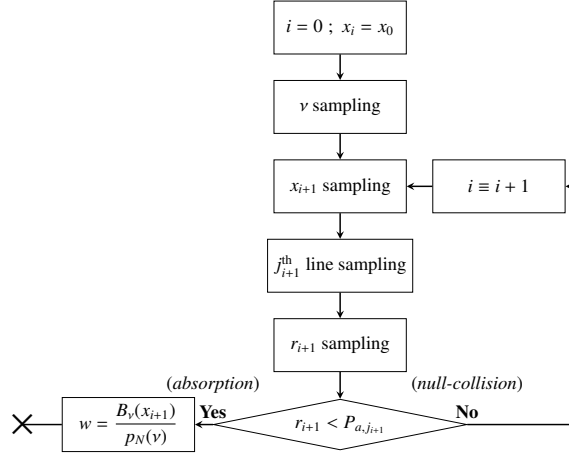
Figure 6: The W-sampling algorithm directly translated from Eq. 16. Same as in Fig. 5 with only an additional sampling of wavenumber v according to the arbitrary probability density $p_N(v)$ and a modification of the final weight w.

3.4. Selection of the free parameters

Again, from a purely theoretical point of view, $\hat{k}_v$, $P_{J_i}(j_i)$ and $p_N(v)$ can be chosen arbitrarily: the expectation of W will always be equal to $I(x_0)$ and provided its variance is finite, the algorithm will converge to the expected value regardless of these choices. They only affect the speed of convergence. But practically speaking, a selection must be made for these free parameters. One possible selection procedure is detailed in Appendix A and we used this procedure for all the simulations results of the present article. This selection resulted from a tedious trial-and-error process and we are still attempting to find theoretical arguments, either explaining why this choice leads to good convergence rates in quite distinct atmospheric and combustion configurations, or justifying a better optimized choice. We believe that the zero-variance framework [21, 22, 23, 24] should be very useful in this context.

As far as the present proposal is concerned, the main points are the following (see Appendix A):

- In standard null-collision algorithms, it is strongly advised that $\hat{k}_v$ be greater than $k_{a,v}$ at all frequencies. This is because absorption versus null-collision is decided with probability $\frac{k_{a,v}}{\hat{k}_v}$ and this ratio must be lower than one. But here, the decision is made with probability $\frac{h_{a,v,j}}{\hat{k}_v P_j}$. So $\hat{k}_v$ cannot be selected independently of P_j: if the transition is sampled with a probability set that does not sufficiently reflect its contribution, this must be compensated by a greater $\hat{k}_v$ value (see the tuning of α in Appendix A).

- Even if $\hat{k}_v$ needs to be artificially increased because of the imperfect adequation of P_j, it is still useful that its spectral shape reflects the shape of $k_{a,v}$. Indeed, if $\hat{k}_v$ was flat, its value would be adjusted to be at least greater than the value of the absorption coefficient at the center of the stronger line and this would imply a tremendous number of null-collisions at line-wing frequencies. So $\hat{k}_v$ needs somehow to "follow" $k_{a,v}$, whereas the main interest of the present work is precisely to avoid the evaluation of $k_{a,v}$. But any rough information about $k_{a,v}$

gives the order of magnitude of the line-center/line-wing ratio and solves this difficulty. So any available $k_{a,v}$ spectrum can be used provided that it roughly corresponds to the same range of pressure, temperature and concentrations, without needing to consider the detailed spectroscopic assumptions (database, truncations, etc). We are thinking of providing a validated $\hat{k}_v$ database, but at the present stage gaseous-radiation specialists can start with their available high-resolution spectra as indicated in the Appendix.

- Ideally, in terms of convergence, P_j should be proportional to $h_{a,v,j}$. The choice made in the Appendix differs from this because of pure computational constraints: memory access (for large spectral databases) and sampling acceleration. This was only a first attempt to find the minimum computation time as a compromise between reducing the number of samples (when P_j is closely adjusted to $h_{a,v,j}$ the variance of W is minimal) and reducing the sampling time itself (when P_j is easy to handle).

4. Test cases

The algorithm is applied to the six benchmark configurations[3] gathered by Andre and Vaillon in [20] (test case 1 to test case 6). These configurations cover a wide variety of practical applications (from furnaces to plume infrared signature) for which radiative transfer computations may be required. They are adapted from the test cases introduced in [25] (cases 1 and 2), [26] (cases 3, 4 and 5) and [27] (case 6). Each of these configurations involves an emitting and absorbing gaseous column at atmospheric pressure defined over $[0, x_0]$. Scattering phenomena are neglected. The $x = 0$ boundary is a black wall at $0K$. From one configuration to the other, only the domain size (the value of x_0) and the piecewise-defined fields of temperature and of molar fraction are modified. These characteristics are described in Table 1.

Unless otherwise specified, the computation was made using Lorentz profiles and a 25cm^{-1} line-wing truncation. Outgoing intensities averaged over several 25cm^{-1} narrowbands are depicted in the plots of Fig. 7 - 9 for test case 2[4] using several databases (HITRAN 2008 [28], two versions of the HITEMP database: HITEMP and HITEMP 2010 [4]), several intensity cutoffs and several distances of line-wing truncation. For validation, a deterministic high-resolution computation (based on line-by-line spectra) using HITEMP 2010 is also depicted. Each simulation used the same $\hat{k}_v$, P_{J_i} and p_N choices: no additional preprocessing was therefore required for each of these spectroscopic-hypothesis changes. Mean intensities were computed in 0.1 to 2.4 seconds for a 1% standard deviation on a single core of an Intel Core i7 processor (2.8GHz). The size of the spectroscopic database, the intensity-cutoff level and the line-truncation distance have little influence on the computation time (it depends mainly on the narrowband itself). These computation times exclude the loading of spectroscopic data into memory and the eventual preprocessing steps

---

[3] Practically, when dealing with polymolecular gases, the molecular species index is first sampled before a transition for this species is sampled. Also to be mentioned: the black wall at $x = 0$ is treated as an infinite absorbing medium; when a sampled location overtakes the wall location, the realization stops.

[4] The particular choice of test case 2 for illustration purposes is fully arbitrary. Similar results were obtained for the other studied configurations.

Case	Size [m]	Columns [m]	Temperature [K]	x_{CO_2} []	x_{H_2O} []	x_{N_2} []
1	1.1	$x \in [0; 0.1]$	1500	0.5	-	0.5
		$x \in [0.1; 1.1]$	500	0.05	-	0.95
2	1.1	$x \in [0; 0.1]$	1500	-	0.5	0.5
		$x \in [0.1; 1.1]$	500	-	0.05	0.95
3	8	$x \in [0; 1.5]$	$400 + 2000x/1.5$	0.1	0.2	0.7
		$x \in [1.5; 8]$	$800 + 1600(8 - x)/6.5$	0.1	0.2	0.7
4	8	$x \in [0; 1.5]$	$400 + 2000x/1.5$	0.1	-	0.9
		$x \in [1.5; 8]$	$800 + 1600(8 - x)/6.5$	0.1	-	0.9
5	8	$x \in [0; 1.5]$	$400 + 2000x/1.5$	-	0.2	0.8
		$x \in [1.5; 8]$	$800 + 1600(8 - x)/6.5$	-	0.2	0.8
6	200.8	$x \in [0; 0.2]$	1500	-	0.15	0.85
		$x \in [0.2; 0.4]$	1200	-	0.12	0.88
		$x \in [0.4; 0.6]$	900	-	0.09	0.91
		$x \in [0.6; 0.8]$	500	-	0.06	0.94
		$x \in [0.8; 200.8]$	300	-	0.03	0.97

Table 1: Dimensions, temperature and molar fractions fields for the six test cases. Fields are piecewise-defined in several columns depending on the case.

(in particular concerning the P_{J_i} computation, detailed in Appendix A, which could last from less than one second to a few minutes for the bigger spectroscopic databases). Times related to loading and precomputation are therefore frequently greater than the computation times themselves. However, once they are done, it is possible to run as many different computations as desired (either to address different observables within the same configuration, or to address various spectral assumptions).
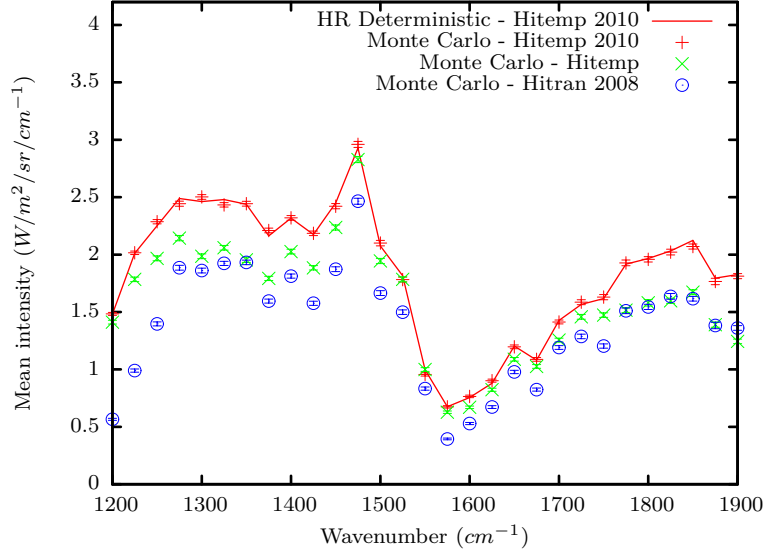

Figure 7: Intensities $I_\nu(x_0)$ averaged on 25cm^{-1} narrowbands computed for test case 2. This computation used the algorithm of Fig. A.10 for several spectroscopic databases: HITEMP 2010, HITEMP and HITRAN 2008. The results are given together with their statistical uncertainties. Each point was obtained using 10^4 independent statistical realizations. A deterministic computation is also presented (solid line) for the HITEMP 2010 database.

Table. 2 gathers intensities integrated over the 10 to 15000 cm^{-1} range, noted $\tilde{I}_{mcm}(x_0)$. They are computed using several spectroscopic databases for test case 1 to test case 6 with 10^6 independent realizations [5]. Also displayed is the computation time $t_{1\%}$ required to get a 1% standard deviation. These results are compared with those obtained with a deterministic high-resolution approach, $\tilde{I}_{hr}(x_0)$, using the exact same line models and data (but requiring the production of line-by-line spectra). They are also compared with those obtained by Andre and Vaillon in [20], here noted $\tilde{I}_{a,v}(x_0)$. The consistency of $\tilde{I}_{mcm}(x_0)$ and $\tilde{I}_{hr}(x_0)$ is full and the differences with $\tilde{I}_{a,v}(x_0)$ are within the observed effects of databases changes. The computation time required by the Monte Carlo algorithm to get a 1% standard deviation is comprised between 0.97 and 9.86s (on a single core of an Intel Core i7 processor - 2.8GHz still excluding loading and precomputation times). Because of the stochastic nature of the method, these computation times (required to integrate $I_\nu(x_0)$ over a large spectral interval) are close to those of Fig. 7-9 (integration over a narrowband): the size of the considered band has little influence.

---

[5]Thanks to the nature of our null-collision algorithm, for the cases 3, 4 and 5, by comparison with [20], a piecewise-constant approximation is no longer needed to handle the linear temperature field.
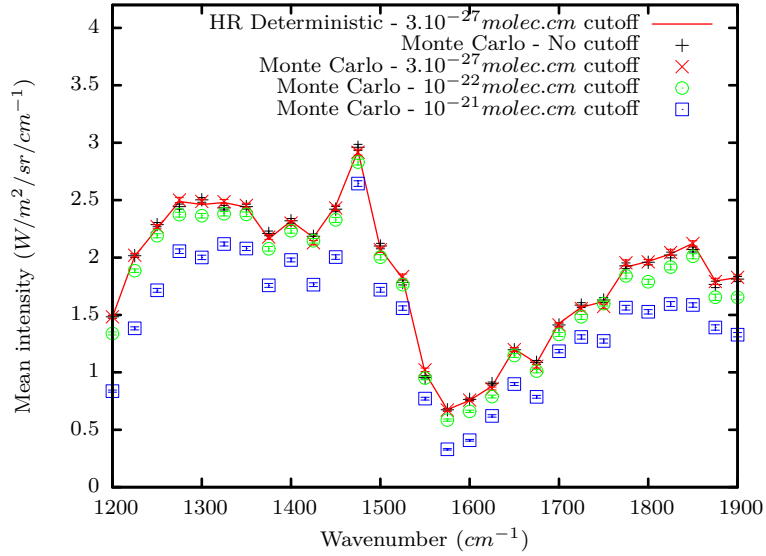
Figure 8: Intensities $I_\nu(x_0)$ averaged on 25cm^{-1} narrowbands computed for test case 2 with the HITEMP 2010 database. This computation used the algorithm of Fig. A.10 for several intensities cutoff at 1500K: 3×10^{-27}, 10^{-22}, 10^{-21} molec.cm and no cutoff. The results are given together with their statistical uncertainties. Each point was obtained using 10^4 independent statistical realizations. A deterministic computation is also presented (solid line) for the HITEMP 2010 database for a 3×10^{-27} molec.cm cutoff at 1500K.
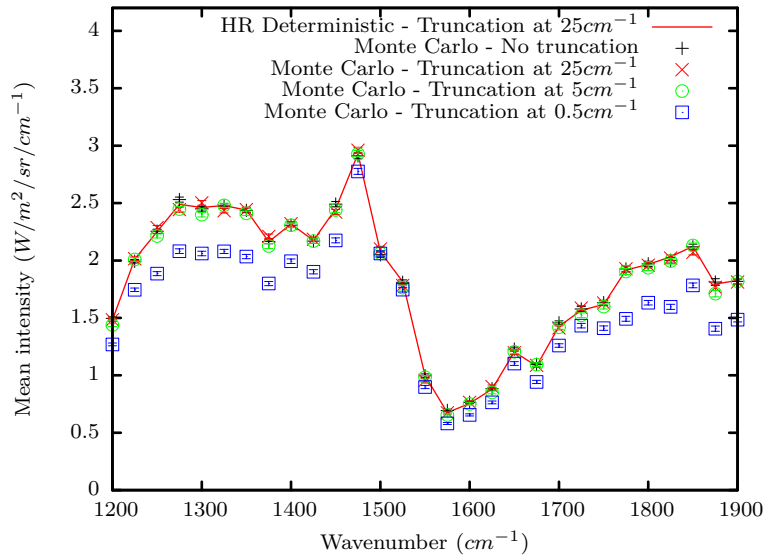

Figure 9: Intensities $I_\nu(x_0)$ averaged on 25cm^{-1} narrowbands computed for test case 2 with the full HITEMP 2010 database. This computation used the algorithm of Fig. A.10 for several distances of line-wing: 25, 5, 0.5cm^{-1} and no truncation. The results are given together with their statistical uncertainties. Each point was obtained using 10^4 independent statistical realizations. A deterministic computation is also presented (solid line) for the HITEMP 2010 database with a 25cm^{-1} truncation.

Case	Databases	Monte Carlo (10^6 realizations)			High Res.	Andre,Vaillon
		$\tilde{I}_{mcm}(x_0)$ (W/m^2/sr)	σ (W/m^2/sr)	$t_{1\%}$ (s)	$\tilde{I}_{hr}(x_0)$ (W/m^2/sr)	$\tilde{I}_{a,v}(x_0)$ (W/m^2/sr)
1	CDSD-1000	3125.61	4.42	0.97	3126.06	3105
	CDSD-4000	3146.25	4.53	1.10	3150.32	
2	HITEMP	3315.11	8.15	1.38	3311.88	4161
	HITEMP 2010	4545.05	9.83	1.11	4558.68	
3	CDSD-1000 & HITEMP	39223.87	51.56	1.75	39202.5	39331
4	CDSD-1000	12325.99	16.16	1.26	12320.1	11956
5	HITEMP	38240.31	49.58	1.27	38215.0	39144
6	HITEMP	885.93	3.93	9.86	886.55	-
	HITEMP 2010	1066.92	4.30	7.39	1069.81	-

Table 2: Intensities integrated over the 10 to 15000 cm^{-1} range for the six test cases introduced in Table 1 and for several databases (CDSD-1000 versus CDSD-4000 for CO_2 and Hitemp versus Hitemp 2010 for H_2O). The estimation $\tilde{I}_{mcm}(x_0)$ obtained with the proposed method are given together with its standard deviation σ for 10^6 realizations. We give the time $t_{1\%}$ required to reach a 1% accuracy, that is related to the simulation time t according to $t_{1\%} = 10^4 t \sigma^2$. They can be compared to those obtained with a deterministic high-resolution computation $\tilde{I}_{hr}(x_0)$ or to the line-by-line values $\tilde{I}_{a,v}(x_0)$ obtained by Andre and Vaillon in [20]. Computation times exclude the loading of spectroscopic data into memory and the precomputations introduced in Appendix A.

All the usual benefits of Monte Carlo methods are apparent here: the computation requirements depend very little on the complexity level (number of retained transitions at each wavenumber, size of the integration range and therefore number of lines with a first-order influence). So we can expect that further complexities will induce small increases in the computation times, mainly when considering real system geometries or multiple-scattering phenomena (at least no specific difficulty is to be expected in comparison to the standard practice [2]).

5. Perspectives outside Monte Carlo applications

Molecular spectroscopic databases [3, 4, 5, 6] are being continuously upgraded by a wide, international and very active community of spectroscopists. Physicists and engineers rely on these databases each time they analyze radiative experiments or address heat transfer questions involving molecular gases. But even for the simplest gases at the coldest temperatures, in most applicative contexts the number of state-transitions that need to be considered is large, or even extremely large when dealing, for instance with water vapor or carbon dioxide on the full infrared or visible spectral ranges. This has always led spectroscopists to envisage statistical representations. The best-known example corresponds to the design of statistical narrowband models that allowed the evaluation of spectrally-averaged transmissivities from weighted averages of line intensities, line widths, etc. But in practice, this direct relationship between line-statistics and band-average transmissivities was rapidly lost. The idea of using statistical models was kept (notably with the continuous use of Malkmus modeling assumptions [29]), but the free parameters were more and more

commonly adjusted from line-by-line calculations: no statistics were made on the databases themselves. They were only used to produce high-resolution absorption-spectra with deterministic approaches, with these spectra being used to compute average-transmissivities, and the narrowband parameters then being adjusted to fit these transmissivities. Later, with the use of k-distributions [30, 31], the same logic was continued: the approach was statistical but it started from high-resolution spectra instead of spectroscopic databases.

As illustrated in the previous section, it is now possible to gather the complexity of radiative transfer and that of molecular-transitions into a single tractable statistical problem. Any radiative transfer observable can be viewed as the expectation of a random variable defined using only transition properties. The information usually carried by the absorption coefficient, in standard statistical approaches to radiative transfer, is here statistically reconstructed by a combination of transitions, via the recursion of null-collisions. Before null-collisions were introduced, such a statistical formulation was impossible because Beer extinction is nonlinear: the absorption coefficient could always be split into transition contributions, but these transitions could not be extracted from the exponential. Null-collisions have achieved this task:

- within the exponential, only $\hat{k}_\nu$ appears (see the definition of $\hat{p}_{X_i}$),

- $k_{a,\nu}$ appears outside, in a linear manner, via the absorption and null-collision probabilities (see the definitions of P_a and P_n),

- because of this linearity, $k_{a,\nu}$ can be replaced by transition statistics,

- the nonlinearity of the exponential is reconstructed by the forward-scattering recursion of null-collisions [1, 32].

We have seen that this has immediate numerical consequences, but we believe that Eq. 14 can also be a starting point for theoretical studies of spectroscopic models under radiative transfer perspectives, as when Malkmus addressed spectrally averaged column-transmissivities. Furthermore, Eq. 14 can be derived with respect to any parameter π of the line-shape model to build a new statistical picture, this time of the sensitivity to π of $I_\nu(x_0)$ [21, 22, 33, 34]:

$$\partial_\pi I_\nu(x_0) = E\left(B_\nu(Y)\sum_{i=1}^{+\infty}\left(A_i\prod_{q=1}^{i-1}(1-A_q)\right)\left(\partial_\pi ln(P_{a,J_i}) + \sum_{m=1}^{i-1}\partial_\pi ln(P_{n,J_m})\right)\right) \qquad (17)$$

This means that for a given radiative-transfer quantity φ of interest, we can always define functions of random transitions with the property that their expectation equals φ or the sensitivity of φ to any spectroscopic parameter. On this basis, we can envisage modeling efforts, inspired by the Malkmus approach, i.e. statistically modeling the transitions and adjusting the free parameters to fit radiative quantities of interest, as well as their sensitivities to thermodynamic state-variables. We believe that this would be particularly meaningful in the context of general circulation modeling (radiative transfer parameterization) when thinking of planetary atmospheres such as Venus, or exoplanets presenting similar radiative-transfer challenges [35, 36, 37]).

6. Conclusion

The introduction of null-collisions allows the design of reference Monte-Carlo algorithms in which absorption properties are addressed during the simulation, via a sampling of molecular transitions. The costly step of producing line-by-line spectra for given sets of thermodynamic properties (pressure, temperature, composition) and spectroscopic hypotheses (database, truncation, intensity cutoff, *etc.*) is no longer necessary. The interpolation of these spectra during the computation also disappears: all the contributions to absorption are computed for the exact wavenumber and the exact local properties. Radiative transfer is now thought out at the scale of molecular transitions.

In order to prove the practical feasibility, it was necessary to make a choice for free parameters that have an important impact on convergence rates. This choice (developed in Appendix A) results from a tedious trial-and-error-process and is only a first attempt that need to be continued. Nevertheless, the first results tested against several distinct one-dimensional and purely absorbing test cases, from atmospheric to combustion and IR plume-detection applications, lead to satisfying results in terms of convergence rate and precision. Important work remains to be done to see if this method and the choice of free parameters can have a general and operational value for any kind of radiation problem or on the contrary should be considered on a case-by-case basis for specific applications.

In particular, specific attention should be given to the validity of the estimated standard deviation. Thus, when exploring the behavior of new Monte Carlo algorithms, when it is quite distinct from previous practices, there is always a risk that some events with important statistical weight are not sampled. Such statistics lead to biased estimation of standard deviation that can no longer be considered as a statistical error. Theoretically speaking, this question of well-behaved statistics is always very difficult because the answer requires the knowledge of higher-order W-statistics. Very commonly, only the practice can help. For standard Monte Carlo algorithms in the context of radiative-transfer, a huge amount of practical experience has been reported and we know for sure that the Monte Carlo weights are well behaved, which means that the computed errorbars can be trusted. This explains why Monte Carlo simulations are accepted as reference simulations. But here the multiple sampling of lines introduces new statistical behaviors and we observed, during the first stages of our developments, that a poor selection of the free parameters could be pathological: the true solution could be outside the computed confidence interval. Even if we are now confident that our selection of free parameters leads to well-behaved statistics for the 7 studied test cases (which have been validated against line-by-line simulation) and that the errorbars will remain correct when using the present set of free parameters for a wide range of radiative transfer applications, we strongly advise the user to implement the variance-of-variance (VOV) computation and systematically check its value before sufficient experience is gathered for a given configuration family. Even if it is not a complete guarantee, the variance of variance is a way to detect such pathological behavior when reference solutions are not available. Its expression is given by:

$$VOV = \frac{S_4 - \frac{4}{M}S_1 S_3 + \frac{8}{M^2}S_2 S_1^2 - \frac{4}{M^3}S_1^4 - \frac{1}{M}S_2^2}{\left(S_2 - \frac{1}{M}S_1^2\right)^2} \tag{18}$$

with $S_k = \sum_{m=1}^{M} w_m^k$. As advised in [38], if the computed VOV does not exceed 0.1, the standard deviation can be

considered as reliable. With the set of free parameters that we finally adopted, we observed that *VOV* was below 0.005 in all test cases (for 10^6 realizations).

Our pure-absorbing test cases include large heterogeneities in temperature, pressure and concentrations, including the strong line-width variations between ground level and high altitudes in our first atmospheric example, and we noted no mechanism increasing the variance as a function of these typical features. Therefore, we cannot think of any configurations where the sampling of lines would introduce specific variances of the Monte Carlo weight. Of course, this does not mean that computation times will not increase when dealing with multiple scattering and complex geometries. This extension is in progress. However, at present, there does not seem to be any particular reason that would indicate that multiple scattering and complex geometries would have a different impact on computation than for previous null-collision Monte Carlo algorithms [2]. In the algorithmic structure of Appendix A, path-tracking including multiple reflection and multiple scattering will only be a refinement of the "x_{j+1} sampling" step: the line-sampling proposal presented here is strictly independent of the optico-geometrics.

Appendix A. Resulting algorithms and $\hat{k}_\nu$, P_{J_i}, p_N choices

Starting from Eq.16, the two modifications proposed in Sec. 3 (an energy-partitioning approach and a modified absorption probability) lead to the algorithm displayed in Fig. A.10. But choices still remain to be made for $p_N(\nu)$, $P_{J_i}(j_i)$, $\hat{k}_\nu$ as well as for the threshold ζ at which the energy-partitioning algorithm switches to a standard Russian-roulette algorithm. ζ is retained as equal to 0.5 in accordance with [2]. $p_N(\nu)$ only plays a role in computation-time enhancement. We first used p_N independent of wavenumbers. This led to very satisfactory computation times. We tried however to enhance the wavenumber sampling procedure by first sampling a narrowband i_b, among N_b bands of width $\Delta \nu = 25 cm^{-1}$, according to the probability

$$P_b(i_b) = \frac{\tilde{I}_b(i_b) + \frac{1}{N_b} \sum_{b=1}^{N_b} \tilde{I}_b(b)}{2 \sum_{b=1}^{N_b} \tilde{I}_b(b)} \tag{A.1}$$

and then sample the wavenumber uniformly within the narrowband. In Eq. A.1, $\tilde{I}_b(i_b)$ corresponds to precomputed approximate solutions using a standard narrowband model (Malkmus model with the Curtis-Godson approximation) over which we applied an offset to compensate underestimations of some of the weaker bands. To reach the same accuracy, the required computation times were reduced by a factor of between 2 and 5. Obviously, such a wavenumber sampling enhancement is therefore only of limited interest. This is certainly due to the fact that we weight the bands, but not the wavenumbers within each band. A clever solution to weight the wavenumbers would be to make use of the approach introduced in [9].

The main difficulty is the choice of P_{J_i} and $\hat{k}_\nu$. They are highly critical in terms of convergence. Indeed, as much as possible, P_{n,J_i} has to be greater than zero in Eq. 13: if not, negative values of the null-collision frequency are handled as advised in [1], but this severely increases the variance. We therefore need to avoid such occurrences which leads to
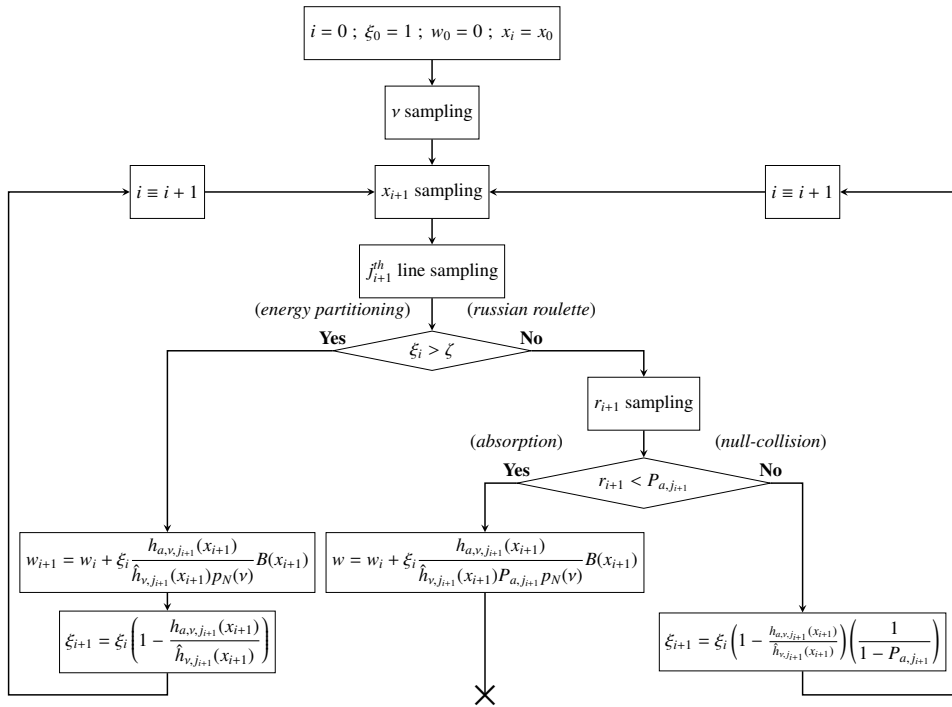
22

Figure A.10: Refined *W*-sampling algorithm. Starting from location x_0 where the intensity is estimated, a wavenumber ν, a location x_1 of collision and a line j_1 are successively sampled. Then, if $\xi_0 > \zeta$, the energy partitioning branch is followed where an absorption contribution is added to the Monte-Carlo weight and where the null-collision term ζ_1 is computed, finally the branch loops on the step of collision-location sampling with an index i incremented. If $\xi_0 \leq \zeta$, the Russian roulette branch is followed, and the type of collision is defined by sampling uniformly a number r_1 between 0 and 1. If $r_1 < P_{a,j_1}(x_1)$ it is an absorption, so the sample w_r of the random variable W is then computed and the loop stops; otherwise it is a null-collision and the algorithm loops at the step of collision-location sampling with an index i incremented.

the following condition:

$$\hat{h}_{v,j_i}(x_i) \geq h_{a,v,j_i}(x_i) \tag{A.2}$$

or equivalently (see Sec. 2.4):

$$\hat{k}_v(x_i) \geq \frac{h_{a,v,j_i}(x_i)}{P_{J_i}(j_i)} \tag{A.3}$$

The P_{J_i} and $\hat{k}_v$ choices, although fully arbitrary in pure theoretical terms, are very much constrained in practice, in a closely related manner. We retained the following strategy

- we chose $\hat{k}_v = \alpha \tilde{k}_v$ where $\tilde{k}_v$ is a precomputed high resolution spectrum and $\alpha > 1$,

- we devoted most of our attention to the P_{J_i} choice,

- we finally adjusted α to ensure Eq. A.3.

Choosing the high resolution spectrum $\tilde{k}$. This is the only stage at which we make use of precomputed high-resolution spectra. The essential point is that these spectra do not need to be in coherence either with the local thermodynamic and composition variables, or with the spectral assumptions made for the Monte Carlo computation. In practice, we used available spectra that were computed in other contexts using CDSD-1000 and HITEMP, and the most intense lines from CDSD-4000 and HITEMP 2010 have been added to these spectra in order to ensure Eq. A.3 for the most critical wavenumbers in all cases (the resulting spectra are distributed within the EDStar environment [39]). These spectra were kept even when testing other spectroscopic-databases, or testing other line-profile and intensity-cutoff assumptions than those used for their production. Any spectrum can be used as α will then be increased to ensure Eq. A.3 for all wavenumbers. The benefit of these precomputed spectra being of high quality is only a reduction of the computation time (lower α values will be required and therefore fewer null-collisions will occur). In test case 1 to test case 6, we split the medium into a few subdomains that roughly correspond to the medium heterogeneity (at all locations within a subdomain, the composition and thermodynamic variables take similar values, typically 2 to 5 subdomains) and for each subdomain we retained one of our precomputed spectra, the one corresponding to the closest thermodynamic state. Again, this choice does not affect the Monte Carlo solution but only the computation speed. α was finally tuned to $\alpha = 50$ when using our high resolution spectra [6]. We also tested the effects of using no available spectrum and taking $\tilde{k}$ independent of wavenumber. Table A.3 displays the corresponding increase in the computation time. Clearly, by comparison with the p_N choice, the $\tilde{k}_v$ choice has a very significant impact: the computation time is increased by a factor of 2000 for test case 6.

Choosing P_{J_i}. This is undeniably the step for which the most significant upgrades can be expected in future works. It is in this sense that the Monte Carlo simulations presented in the present article are only a first feasibility test. Any

---

[6]Even with $\alpha = 50$, some rare events lead to a violation of A.3. We therefore handle them as advised in [1] and increase $\tilde{k}_v$ for the following sampling events at the same wavenumber.

Test case	$t_{1\%}$ for $\hat{k}$ uniform (s)	Databases	Factor of comp. time increase
Case 1	57.6	CDSD-1000	59.4
Case 2	215.0	HITEMP	155.8
Case 3	457.2	CDSD-1000 & HITEMP	261.3
Case 4	365.7	CDSD-1000	290.2
Case 5	355.9	HITEMP	280.2
Case 6	19825.2	HITEMP	2010.7

Table A.3: Computation times $t_{1\%}$ to reach a 1% standard deviation for the six cases with $\hat{k} = \alpha \tilde{k}$ uniform ($\hat{k} = 30cm^{-1}$ for test cases 1 to 5 and $\hat{k} = 10cm^{-1}$ for test case 6). Also given are the employed spectroscopic databases and the increase factor of computation time associated with $\hat{k}$ uniform by comparison with the $\hat{k}_\nu$ choice of Appendix A (where the free parameter α is set as 50).

upgrade of the P_{J_i} choice will be immediately translated into lower α values. The fact that $\alpha = 50$ is required (which we observed was the case even in cases where $\tilde{k}_\nu$ was very accurate) indicates that our present choice should be highly improvable. However, one point is already theoretically quite clear: we used zero-variance reasoning to establish that the ideal P_{J_i} choice should be $P_{J_i}^{\text{ideal}}(j_i) = h_{a,\nu,j_i}(x_i)/k_{a,\nu}(x_i)$. Obviously, this cannot be retained in practice because $k_{a,\nu}(x_i)$ is unknown. So the whole question is to choose $P_{J_i}(j_i)$ so that it is close enough to $P_{J_i}^{\text{ideal}}(j_i)$ but remains easy to evaluate (short computation time and limited precomputation requirements). In practice, we define $[\nu_{db,min}, \nu_{db,max}]$ as the interval containing the lines centers corresponding to all transitions within the considered database. We divide this interval into regular bands of width $\delta\nu = 0.3cm^{-1}$. Then we take $P_{J_i}(j_i) = T_U(u) \times Q_{J_i}(j_i, u)$ where

- $T_U(u) \equiv T_{U,\nu}(u, \gamma_Q)$ is a probability associated with the band containing the center of line j_i (indexed u) among all bands:

$$T_U(u) = \frac{\text{atan}\left(\frac{b(u)-\nu}{\gamma_Q}\right) - \text{atan}\left(\frac{a(u)-\nu}{\gamma_Q}\right)}{\text{atan}\left(\frac{\nu_{db,max}-\nu}{\gamma_Q}\right) - \text{atan}\left(\frac{\nu_{db,min}-\nu}{\gamma_Q}\right)} \tag{A.4}$$

where γ_Q is a free parameter and $[a(u), b(u)]$ is the spectral interval of the uth band: $b(u) - a(u) = \delta\nu$. With such a definition, sampling a band index according to T_U is equivalent to sampling a wavenumber according to a Lorentz profile of width γ_Q around ν and retaining the corresponding band (see step 1 of Fig. A.11)[7]. The efficiency of the algorithm can be affected by the arbitrary choice of γ_Q parameter. The best convergence rates were obtained here using $\gamma_Q = 1cm^{-1}$ for temperatures $T < 500K$, $\gamma_Q = 0.5cm^{-1}$ for $T \in [500K, 1000K]$, $\gamma_Q = 0.25cm^{-1}$ for $T \in [1000K, 1500K]$ and $\gamma_Q = 0.1cm^{-1}$ for $T > 1500K$. This temperature-dependent definition accounts for the broadening/narrowing of "real" line half-widths with regard to temperature. For the atmospheric configuration treated in Sec. 1, where smaller gradients of temperature are encountered, γ_Q was defined as a function of pressure, which plays, in this case, the biggest role in line half-widths: $\gamma_Q = \frac{P}{P_{ref}} 1cm^{-1}$ with $P_{ref} = 1atm$, which is consistent with the values used for the plume detection cases.

---

[7]If line-wings are truncated at a distance $\delta\nu_{trunc}$, then $[\nu_{db,min}, \nu_{db,max}]$ is replaced with $[\nu - \delta\nu_{trunc}, \nu + \delta\nu_{trunc}]$.

- $Q_{J_i}(j_i, u) \equiv Q_{J_i, \nu}(j_i, u)$ is a probability associated with line j_i among all the lines centered within the same band (this probability simply reflects the line intensity):

$$Q_{J_i}(j_i, u) = \frac{S_{j_i}}{\sum_{q=1}^{N_t(u)} S_{q_i}} \tag{A.5}$$

where S_{j_i} is the intensity of line j_i at a given temperature T and $N_t(u)$ is the number of line-centers within band u. Q_{J_i} is precomputed for each of the 2 to 5 subdomains already defined for the $\tilde{k}_\nu$ definition. This precomputation requires between less than 1 second (for HITRAN database) to 5 minutes (for CDSD-4000 database) for each subdomain on a single core of an Intel Core i7 processor (2.8GHz). However, this precomputation can be stored once for all, for a given temperature and a given database.

With this choice, the lines that are preferentially sampled are the most intense ones, centered near the wavenumber of interest and only the inexpensive T_U computations are performed on the fly. The resulting P_{J_i} sampling procedure is illustrated in Fig. A.11.
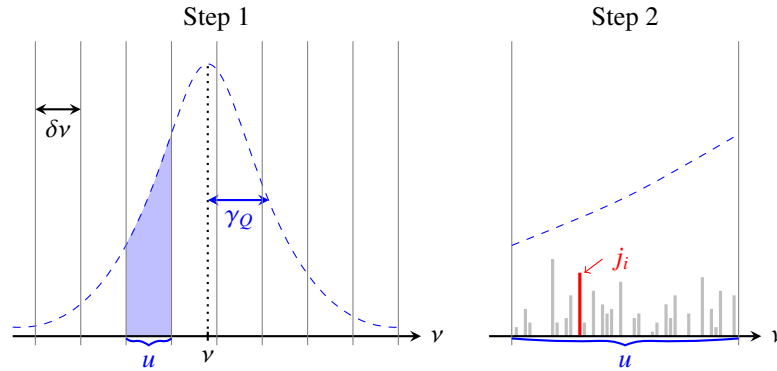

Figure A.11: Sampling procedure of P_{J_i}. *Step 1:* a band u (of width $\delta\nu$) is determined by sampling a wavenumber according to a Lorentz profile of width γ_Q centered in ν. *Step 2:* a line j_i is sampled among the $N_t(u)$ lines centered in u according to its line intensity.

[1] Galtier, M., Blanco, S., Caliot, C., Coustet, C., Dauchet, J., El Hafi, M., et al. Integral formulation of null-collision Monte Carlo algorithms. Journal of Quantitative Spectroscopy and Radiative Transfer 2013;125:57–68.

[2] Eymet, V., Poitou, D., Galtier, M., El Hafi, M., Terrée, G., Fournier, R.. Null-collision meshless Monte-Carlo - Application to the validation of fast radiative transfer solvers embedded in combustion simulators. Journal of Quantitative Spectroscopy and Radiative Transfer 2013;129:145–157.

[3] Jacquinet-Husson, N., Crepeau, L., Armante, R., Boutammine, C., Chédin, a., Scott, N., et al. The 2009 edition of the GEISA spectroscopic database. Journal of Quantitative Spectroscopy and Radiative Transfer 2011;112(15):2395–2445.

[4] Rothman, L., Gordon, I., Barber, R., Dothe, H., Gamache, R., Goldman, A., et al. HITEMP, the high-temperature molecular spectroscopic database. Journal of Quantitative Spectroscopy and Radiative Transfer 2010;111(15):2139–2150.

[5] Rothman, L., Gordon, I., Babikov, Y., Barbe, a., Chris Benner, D., Bernath, P., et al. The HITRAN2012 molecular spectroscopic database. Journal of Quantitative Spectroscopy and Radiative Transfer 2013;130:4–50.

[6] Tashkun, S., Perevalov, V.. CDSD-4000: High-resolution, high-temperature carbon dioxide spectroscopic databank. Journal of Quantitative Spectroscopy and Radiative Transfer 2011;112(9):1403–1410.

[7] Modest, M.. the Monte Carlo Method Applied To Gases With Spectral Line Structure. Numerical Heat Transfer, Part B: Fundamentals 1992;22(3):273–284.

[8] Surzhikov, S.T.. The Use of Monte Carlo Simulation Methods to Calculate the Radiation of Jets of Combustion Products in View of Rotational Spectral Structure. High Temperature - Translated from Teplofizika Vysokikh Temperatur 2003;41(5):694–707.

[9] Feldick, A., Modest, M.F.. An improved wavelength selection scheme for Monte Carlo solvers applied to hypersonic plasmas. Journal of Quantitative Spectroscopy and Radiative Transfer 2011;112(8):1394–1401.

[10] Ren, T., Modest, M.F.. A Hybrid Wavenumber Selection Scheme for Line-By-Line Photon Monte Carlo Simulations in High-Temperature Gases. Journal of Heat Transfer 2013;135(8):084501.

[11] Fomin, B.A.. Monte-Carlo algorithm for line-by-line calculations of thermal radiation in multiple scattering layered atmospheres. Journal of Quantitative Spectroscopy and Radiative Transfer 2006;98(1):107–115.

[12] Ewen, G.B.L., Grainger, R.G., Lambert, a., Baran, a.J.. Infrared radiative transfer modelling in a 3D scattering cloudy atmosphere: Application to limb sounding measurements of cirrus. Journal of Quantitative Spectroscopy and Radiative Transfer 2005;96(1):45–74.

[13] Fomin, B., Falaleeva, V.. A polarized atmospheric radiative transfer model for calculations of spectra of the stokes parameters of shortwave radiation based on the line-by-line and monte carlo methods. Atmosphere 2012;3(4):451–467.

[14] Madhusudhan, N., Knutson, H., Fortney, J.J., Barman, T.. Exoplanetary Atmospheres. 2014.

[15] Allard, F., Homeier, D., Freytag, B.. Models of very-low-mass stars , brown dwarfs and exoplanets. Philosophical Transactions of the Royal Society A 2012;370:2765–2777.

[16] Rajpurohit, A.S., Reylé, C., Allard, F., Scholz, R.D., D., H., Schultheis, M.. High-resolution spectroscopic atlas of M subdwarfs. Effective temperature and metallicity. Astronomy and Astrophysics 2014;564:A90.

[17] Husser, T.o., Berg, S.W.V., Dreizler, S., Homeier, D., Reiners, A., Hauschildt, P.H.. A new extensive library of PHOENIX stellar atmospheres and synthetic spectra. Astronomy and Astrophysics 2013;553:A6.

[18] LUPM, . The POLLUX Database of Stellar Spectra. 2015. URL: `http://npollux.lupm.univ-montp2.fr/`.

[19] Ellingson, R.G., Ellis, J., Fels, S.. The intercomparison of radiation codes used in climate models: Long wave results. Journal of Geophysical Research 1991;96(D5):8929.

[20] André, F., Vaillon, R.. A nonuniform narrow band correlated-k approximation using the k-moment method. Journal of Quantitative Spectroscopy and Radiative Transfer 2010;111(12-13):1900–1911.

[21] Dauchet, J., Blanco, S., Cornet, J.F., El Hafi, M., Eymet, V., Fournier, R.. The practice of recent radiative transfer Monte Carlo advances and its contribution to the field of microorganisms cultivation in photobioreactors. Journal of Quantitative Spectroscopy and Radiative Transfer 2012;128.

[22] De La Torre, J., Baud, G., Bézian, J.J., Blanco, S., Caliot, C., Cornet, J.F., et al. Monte Carlo advances and concentrated solar applications. Solar Energy 2013;103:653–681.

[23] Assaraf, R., Caffarel, M.. Zero-variance principle for monte carlo algorithms. Physical review letters 1999;83(23):4682–4685.

[24] Hoogenboom, J.E.. Zero-Variance Monte Carlo Schemes Revisited. Nuclear science and Engineering 2008;160:1–22.

[25] Soufiani, A., Hartmann, J.M., Taine, J.. Validity of band-model calculations for CO2 and H2O applied to radiative properties and conductive-radiative transfer. Journal of Quantitative Spectroscopy and Radiative Transfer 1985;33(3):243–257.

[26] Liu, F., Smallwood, G.J., Gulder, O.. Application of the statistical narrow-band correlated-k method to non-grey gas radiation in CO2-H2O mixtures : approximate treatments of overlapping bands. Journal of Quantitative Spectroscopy and Radiative Transfer 2001;68:401–417.

[27] Rivière, P., Soufiani, A., Taine, J.. Correlated-k and fictitious gas methods for H2O near 2.7 [mu] m. Journal of Quantitative Spectroscopy and Radiative Transfer 1992;48(2):187–203.

[28] Rothman, L., Gordon, I., Barbe, A., Benner, D.C., Bernath, P., Birk, M., et al. The HITRAN 2008 molecular spectroscopic database. Journal of Quantitative Spectroscopy and Radiative Transfer 2009;110(9-10):533–572.

[29] Malkmus, W.. Random Lorentz Band Model with Exponential-Tailed S-1 Line-Intensity Distribution Function. Journal of the Optical Society of America 1967;57(3):323–329.

[30] Taine, J., Soufiani, A.. Gas IR radiative properties: from spectroscopic data to approximate models. Advances in heat transfer 1999;33:295–414.

[31] Lacis, A., Oinas, V.. A description of the correlated k distribution method for modeling nongray gaseous absorption, thermal emission, and multiple Scattering in Vertically Inhomogeneous Atmospheres. Journal of Geophysical Research: Atmospheres (1984–2012) 1991;96.

[32] Longo, S.. Direct derivation of Skullerud's Monte Carlo method for charged particle transport from the linear Boltzmann equation. Physica A: Statistical Mechanics and its Applications 2002;313(3-4):389–396.

[33] Lataillade, A.D.G.D., Blanco, S., Clergent, Y., Dufresne, J.l., El Hafi, M.. Monte Carlo method and sensitivity estimations. Journal of Quantitative Spectroscopy and Radiative Transfer 2002;75:529–538.

[34] Roger, M., Blanco, S., El Hafi, M., Fournier, R.. Monte Carlo Estimates of Domain-Deformation Sensitivities. Physical Review Letters 2005;95(18):180601.

[35] Lewis, S., Collins, M., Read, P., Forget, F., Hourdin, F., Fournier, R., et al. A climate database for mars. Journal of Geophysical Research 1999;104(E10):24177–24194.

[36] Eymet, V., Fournier, R., Dufresne, J.l., Lebonnois, S., Hourdin, F., Bullock, M.A.. Net exchange parameterization of thermal infrared radiative transfer in Venus' atmosphere. Journal of Geophysical Research 2009;114(E11):E11008.

[37] Lebonnois, S., Hourdin, F., Eymet, V., Crespin, A., Fournier, R., Forget, F.. Superrotation of Venus' atmosphere analyzed with a full general circulation model. Journal of Geophysical Research 2010;115(E6):E06006.

[38] X-5 Monte Carlo Team, . MCNP — A General Monte Carlo N-Particle Transport Code, Version 5, Volume 1: Overview and Theory. Los Alamos; 2008.

[39] STARWest, . EDStaR (STARWest Development Environment). 2015. URL: http://edstar.lmd.jussieu.fr/.