



HAL
open science

Bridging physics and statistical learning methodologies for the accurate modeling of the radiative properties of non-uniform atmospheric paths

F. André, C. Delage, Loic Guilmard, Mathieu Galtier, Céline Cornet

► **To cite this version:**

F. André, C. Delage, Loic Guilmard, Mathieu Galtier, Céline Cornet. Bridging physics and statistical learning methodologies for the accurate modeling of the radiative properties of non-uniform atmospheric paths. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 2024, 320, pp.108961. 10.1016/j.jqsrt.2024.108961 . hal-04740722

HAL Id: hal-04740722

<https://hal.science/hal-04740722v1>

Submitted on 21 Oct 2024

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.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21

**BRIDGING PHYSICS AND STATISTICAL LEARNING METHODOLOGIES
FOR THE ACCURATE MODELING OF THE RADIATIVE PROPERTIES
OF NON-UNIFORM ATMOSPHERIC PATHS**

F. André^{1,2*}, C. Delage¹, L. Guilnard¹, M. Galtier¹, C. Cornet²

¹Univ. Lyon, CNRS, UMR 5008 – CETHIL, Lyon, France

²Univ. Lille, CNRS, UMR 8518 – LOA, Lille, France

* Corresponding Author: frederic.andre@insa-lyon.fr, frederic.andre@univ-lille.fr.

22 ABSTRACT. The objective of the present work is to describe a technique to approximate atmospheric
23 path transmissivities using a recurrent structure, following a method proposed recently but limited to
24 date to high temperature applications. The physical model together with its underlying statistical
25 assumptions is detailed. It is found to involve a rather simple analytical formula that applies both to
26 two-layers systems and to more general multi-layers non-uniform configurations. This treatment of
27 path non-uniformities uses several unknown parameters that are first trained on LBL reference data in
28 two-layers configurations to illustrate the relevance of the proposed approximate model. Then, in a
29 second time, model's parameters are trained on non-uniform path transmission curves representative
30 of multi-layers atmospheres. The corresponding recurrent formulation is shown to provide accurate
31 estimates of transmissivities of non-uniform atmospheric paths (maximum relative errors are below
32 0.35% in all the considered test cases) at a very low CPU cost.

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

1. INTRODUCTION

48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73

Satellite remote sensing is a widely spread method to characterize the state of the atmosphere and to monitor atmospheric components such as water vapor, clouds and aerosols. Atmospheric instruments measure radiances averaged over more or less narrow spectral bands. The most recent generations of satellites produce large amounts of data that need to be converted into geophysical parameters. This requires the developments of fast though accurate numerical methods to infer profiles of temperature, species concentrations, *etc*, from satellite measurements. In most cases, when modeling a radiative intensity in an atmospheric configuration, one needs to account for the contribution of the absorption and / or emission of radiative energy by the gaseous components of the atmosphere. Due to the large amount of data to analyze, the use of high-resolution Line-By-Line (LBL) calculations cannot be considered for this purpose, even if the amount of computational power available has increased significantly in the past years.

Gas radiation modeling has always been intimately connected with statistical sciences [1]. Statistical Narrow Band (SNB) models, for instance, are founded on assumptions about the locations of line centers (randomly spread over the spectral interval) and about the distribution of linestrengths. Combined with some realistic additional hypothesis (spectral lines are assumed entirely contained in the narrow spectral interval, for instance), SNB model theory allows constructing simple though accurate analytical models based on a few parameters (two, for example, in the case of the widely spread SNB model for Lorentz lines based on Malkmus' distribution of linestrengths [2]). k -distribution methods are constructed, both over narrow spectral intervals [1] or over the full spectrum [3], with the help of the distribution function of absorption coefficients over the spectral interval. In the Correlated K-Distribution (CKD) method, statistical assumptions about the relationship between spectra in distinct thermophysical states are made to extend the method from uniform to non-uniform configurations: the so-called "correlation" assumption [1,3].

74 In this context, it is not surprising to see the appearance, in the recent years, of new strategies based on
75 statistical learning methods [4] for fast though accurate modeling of non-uniform gaseous paths.
76 Principle Component Analysis (PCA) based models [5,6] or techniques founded on unsupervised
77 learning / cluster analysis [7] are two examples of such approaches. More recently, several attempts
78 have been reported in the literature to apply methods based on neural networks to model band averaged
79 gas transmittance [8,9] or to treat more general radiative transfer problems [10].

80

81 One of them, Ref. [9], proposes to apply Recurrent Neural Network (RNN) structures to model non-
82 uniform path transmissivities. The main originalities and advantages of this approach compared to brute
83 force machine learning methodologies are:

84 1/ to suggest an RNN-like model mostly built on physical arguments. Indeed, the method consists of
85 the reformulation of a physical model into a recurrent scheme using statistical assumptions on the
86 relationship between gas spectra in distinct states. The resulting network topology avoids the black box
87 disadvantage of most brute force Neural Network models, as all parameters have a clear physical
88 meaning. The proof of concept was given in Ref. [9]. This reference was however restricted to high
89 temperature applications.

90 2/ to propose the use of Exponential Linear Units - ELU [11] - as activation functions. It is well
91 known [12] that the choice of activation functions can have a significant impact both on the
92 computational effort required for the training process and on the extrapolation capability of Neural
93 Network models. The selection of activation functions based on physical arguments is sometimes
94 recommended for use in Physics Informed Neural Networks (PINN) [13]. This strategy consists of
95 selecting activation functions closely related to known solutions of the governing Ordinary / Partial
96 Differential Equations of the physical problem studied instead of generic ones. This approach cannot
97 be applied directly in the case of gas radiation modeling because the problem of path non-
98 uniformities cannot be fully formulated, to the best of our knowledge, using ODEs / PDEs, even if
99 some physical constraints can be expressed in terms of derivatives of the transmissivity (see

100 Appendix III for details). The strategy proposed in Ref. [9], and explored further in the present
101 paper, mostly consists of finding the general form of the solution of implicit equations to treat non-
102 uniform paths (based on the quasi-scaling approximation described in section 2.2) and then use this
103 functional form to select appropriate activation functions. The proposed method thus replaces
104 solutions of ODEs / PDEs in PINN by implicit equations, but the main logic to incorporate physical
105 constraints in a statistical learning framework is the same in both approaches. As will be seen later
106 in this paper, ELUs are then natural choices.

107 3/ to embed physical constraints directly inside the structure of the recurrent network. This can be
108 considered as a type of mathematical induction approach, following the classification proposed in
109 Ref. [14]. Indeed, constraints on the derivative of the transmissivity are directly included in the
110 structure of the model without modifications of the loss function, as often done in Machine Learning
111 studies (called learning bias in Ref. [14]): only the raw data (set of non-uniform path
112 transmissivities) are used in the training process. This fact is discussed in Appendix III.

113
114 The application of recurrent networks to model non-uniform path transmissivities is explored in the
115 present work by treating the general case instead of the simplified model of Ref. [9]. For this purpose,
116 a generalization of the theoretical foundation of the model detailed in [9] is first proposed in section 2.
117 Both the physical model and the statistical assumptions needed to construct the recurrent approximation
118 are described. This section is rather technical but all the necessary details to fully understand the method
119 and follow the steps of the derivation are given. In section 3, a validation of the results of section 2 in
120 cases of two-layers systems (couples of successive layers along a non-uniform atmosphere) is first
121 provided to illustrate the relevance of the proposed approximation. Then, in a second time, the
122 functional form of section 2 is used as a prior knowledge to model non-uniform path transmissivities
123 in multi-layers configurations, using the same recurrent scheme as in Ref. [9]. The role of the functional
124 form of section 2 (that takes the form of a Lévy-Khintchine formula) combined with the recurrent
125 scheme is thus, in a statistical learning context, to reduce the size of the space of possible solutions with

126 the aim of both simplifying the training process and increasing the quality of the transmissivity model.
127 Several application cases are studied in which the present method is assessed against reference LBL
128 calculations for realistic atmospheres.

129

130 Many elements used in the present paper are closely related to results already discussed within the
131 development of the ℓ -distribution approach in Refs. [15-17]. However, no detailed knowledge about
132 this modeling strategy is required to fully understand the present developments, even if some of them
133 use exactly the same mathematical formalism for convenience. When needed, the elements of ℓ -
134 distribution modeling required to understand the present paper are reminded.

135

136

2. GENERAL THEORY

137

138

139 Let us consider a non-uniform atmosphere divided into n uniform layers. A path in the atmosphere
 140 can be represented by a sequence of lengths L_1, \dots, L_n where each L_i , $i = 1, \dots, n$ is the total distance
 141 traveled by the radiative energy in the i -th layer where the thermophysical state of the gas has index
 142 i and where its spectral absorption coefficient at wavenumber ν is $\kappa_{\nu,i}$. The band averaged
 143 transmissivity of the total path represents the fraction of radiative intensity, averaged over the
 144 spectral interval written $\Delta\nu$, that travels in the atmosphere without being absorbed. It is given by:

145

$$146 \quad \tau_{1..n}^{\Delta\nu}(L_1, \dots, L_n) = \frac{1}{\Delta\nu} \cdot \int_{\Delta\nu} \exp(-\kappa_{\nu,1} \cdot L_1 - \dots - \kappa_{\nu,n} \cdot L_n) d\nu \quad (1)$$

147

148 In the present work, as well as in the ℓ -distribution approach, we seek an approximation of this non-
 149 uniform path transmissivity as:

150

$$151 \quad \tau_{1..n}^{\Delta\nu}(L_1, \dots, L_n) = \tau_1^{\Delta\nu}(L_{1..n}) \quad (2)$$

152

153 where $L_{1..n}$ is an effective absorption length defined in such a way that the transmissivity of the non-
 154 uniform path is the same as the transmissivity of an equivalent path of length $L_{1..n}$ in the gas in state
 155 1. Values of $L_{1..n}$ depend on the n lengths L_1, \dots, L_n .

156 In the ℓ -distribution approach, the effective absorption length $L_{1..n}$ is constructed through a recurrent
 157 process defined by the following set of relationships (identical to Eq. (23) of Ref. [9]):

158

$$159 \quad \begin{cases} L_{nn} = L_n \\ L_{i..n} = L_i + \ell_i \circ \tau_{i+1}^{\Delta\nu}(L_{i+1..n}) \end{cases} \quad (3)$$

160 In this method, the treatment of non-uniformities thus starts from one side of the radiation path,
 161 where the state of the gas has index n , and propagates (index i varies from n to 1) until state 1. The
 162 process involves the set of functions ℓ_i , $i = 1, \dots, N$ defined as the inverses of the band averaged
 163 transmissivities $\tau_i^{\Delta\nu}$. These inverse functions obviously exist because band averaged
 164 transmissivities are strictly decreasing functions of the geometrical lengths. The notation “ \circ ” is
 165 used to represent the functional composition, *i.e.* $\ell_i \circ \tau_{i+1}^{\Delta\nu}(L) = \ell_i \left[\tau_{i+1}^{\Delta\nu}(L) \right]$.

166 The physical interpretation of the recurrent scheme is relatively straightforward. Indeed, it consists
 167 of an iterative homogenization process as explained in depth in Ref. [15]. Interested readers should
 168 consult this reference for more details.

169
 170 Application of the recurrent method set by Eq. (3) to generate $L_{1..n}$ requires evaluating the
 171 “coupling” terms involved at each step of the propagative scheme, *i.e.*, the set of functions
 172 $\ell_i \circ \tau_{i+1}^{\Delta\nu}(L)$ over the full collection of layers. We derive in this section an analytical formula to calculate
 173 these functions. In order to simplify the notations, we focus on $\ell_1 \circ \tau_2^{\Delta\nu}(L)$ only (coupling between
 174 states 1 and 2) but the same functional form applies to any other couple of states i and $i+1$. Several
 175 preliminary results are needed for the derivation. They are given in subsection 2.1 (physical model for
 176 $-\partial \ln \tau_1^{\Delta\nu}(L)/\partial L$) and 2.2 (statistical method used to represent the relationship between gas spectra in
 177 distinct states). Section 2.3 and 2.4 provide the derivation of the coupling functions, first in a simplified
 178 case (2.3) and then in the general situation (2.4).

179
 180 **2.1. Physical model for $-\partial \ln \tau_1^{\Delta\nu}(L)/\partial L$**

181 The derivation described in section 2.3, that focuses on the development of an approximation of the
 182 quantity $\ell_1 \circ \tau_2^{\Delta\nu}(L)$ of Eq. (3), has strong connections with mathematical properties of

183 $-\partial \ln \tau_1^{\Delta\nu}(L)/\partial L$. The present section is devoted to the building of an approximate model for this
 184 quantity from which some important and useful properties can be derived.

185

186 We consider a gas spectrum in state 1 made of N overlapping spectral lines. We approximate the
 187 transmissivity of the array of N lines as the product of the transmissivities of the N spectral lines
 188 considered separately, as done in Statistical Narrow Band model theory [1]. We thus consider that
 189 absorption by a single spectral line is statistically independent from all other spectral lines. This
 190 provides:

191

$$192 \quad \tau_1^{\Delta\nu}(L) = \frac{1}{\Delta\nu} \cdot \int_{\Delta\nu} \exp\left(-L \cdot \sum_{i=1}^N \kappa_{\nu,1,i}\right) d\nu = \prod_{i=1}^N \tau_{1,i}^{\Delta\nu}(L) \quad (4)$$

193

194 where:

195

$$196 \quad \tau_{1,i}^{\Delta\nu}(L) = \frac{1}{\Delta\nu} \cdot \int_{\Delta\nu} \exp(-L \cdot \kappa_{\nu,1,i}) d\nu \quad (5)$$

197

198 is the band averaged transmissivity of the i -th spectral line of the gas in the thermophysical state of
 199 layer 1 for a gas path length L .

200 Let ε be a small and positive real number. We define a “useful” range of lengths, written

201 $[L_{\min}(\varepsilon), L_{\max}(\varepsilon)]$, as the set of lengths such that the value of $\tau_1^{\Delta\nu}(L)$ lies in the interval $[\varepsilon, 1-\varepsilon]$,

202 where most of the dynamics of the transmission curve occurs. This means that the minimum and

203 maximum lengths are defined as solutions of the implicit equations:

204

$$205 \quad \begin{cases} \tau_1^{\Delta\nu}[L_{\min}(\varepsilon)] = 1 - \varepsilon \\ \tau_1^{\Delta\nu}[L_{\max}(\varepsilon)] = \varepsilon \end{cases} \quad (6)$$

206 Over $[L_{\min}(\varepsilon), L_{\max}(\varepsilon)]$, the array of N spectral lines absorbs a significant quantity of radiation but
 207 each single line alone only absorbs a small amount of radiation: band averaged transmissivities of single
 208 lines are thus close to 1. This is because, if we assume for simplification that all spectral lines are
 209 identical and statistically independent, the transmissivity of a single line over $[L_{\min}(\varepsilon), L_{\max}(\varepsilon)]$ takes
 210 a value higher than $\varepsilon^{1/N}$ which is thus close to 1 for large N (for instance, values $\varepsilon=0.001$ and $N =$
 211 1,000 provide $\varepsilon^{1/N} \approx 0.993$ – this means that if we have 1,000 identical lines such that the transmissivity
 212 of the set of lines is 0.001, then each spectral line alone has a band averaged transmissivity of 0.993,
 213 which is clearly close to 1). This allows writing:

214

$$215 \quad -\ln \tau_1^{\Delta\nu}(L) = -\ln \left[\prod_{i=1}^N \tau_{1,i}^{\Delta\nu}(L) \right] = -\sum_{i=1}^N \ln \tau_{1,i}^{\Delta\nu}(L) \approx \sum_{i=1}^N [1 - \tau_{1,i}^{\Delta\nu}(L)] \quad (7)$$

216

217 Taking the derivative of the previous equation with respect to the gas path length L then provides the
 218 following approximation:

219

$$220 \quad -\frac{\partial \ln \tau_1^{\Delta\nu}(L)}{\partial L} \approx \sum_{i=1}^N \left[-\frac{\partial \tau_{1,i}^{\Delta\nu}(L)}{\partial L} \right] \quad (8)$$

221

222 As shown in Appendix I, $-\partial \ln \tau_1^{\Delta\nu}(L)/\partial L$ can then be approximated as:

223

$$224 \quad -\partial \ln \tau_1^{\Delta\nu}(L)/\partial L \approx k_{p,1} \cdot \int_0^1 \exp[-k^*(\xi) \cdot L] d\xi \quad (9)$$

225

226 where $k^*(\xi)$, $\xi \in [0,1]$ plays the role of an absorption rate (the unit of k^* is cm^{-1}) that governs the
 227 behavior of $-1/k_{p,1} \cdot \partial \ln \tau_1^{\Delta\nu}(L)/\partial L$ with respect to the gas path length L .
 228 Following Eq. (9), $-1/k_{p,1} \cdot \partial \ln \tau_1^{\Delta\nu}(L)/\partial L$ is thus a Laplace transform.

229
 230 In a neighborhood of a length L_0 chosen in the range $[L_{\min}(\varepsilon), L_{\max}(\varepsilon)]$, application of the mean
 231 value theorem to the integral of Eq. (9) ensures the existence of a real $X(L_0) \in [0,1]$ such that:

$$232 \quad -\frac{1}{k_{p,1}} \frac{\partial \ln \tau_1^{\Delta\nu}(L_0)}{\partial L} = \exp(-k^*[X(L_0)] \cdot L_0) \quad (10)$$

234
 235 For a small positive increment δL_0 , using again the mean value theorem and limiting the analysis
 236 to a first order expansion in δL_0 , we receive the following approximation:

$$237 \quad -\frac{1}{k_{p,1}} \frac{\partial \ln \tau_1^{\Delta\nu}(L_0 + \delta L_0)}{\partial L} = \exp(-k^*[X(L_0 + \delta L_0)] \cdot (L_0 + \delta L_0)) \quad (11)$$

$$238 \quad \approx -\frac{1}{k_{p,1}} \frac{\partial \ln \tau_1^{\Delta\nu}(L_0)}{\partial L} \cdot \exp[-s_1(L_0) \cdot \delta L_0]$$

239
 240 where, by reorganizing Eq. (11) and taking the limit $\delta L_0 \rightarrow 0$:

$$241 \quad s_1(L_0) = \frac{\partial \ln}{\partial L} \left[-\frac{1}{k_{p,1}} \frac{\partial \ln \tau_1^{\Delta\nu}(L=L_0)}{\partial L} \right] = -\frac{\partial^2 \ln \tau_1^{\Delta\nu}(L_0)/\partial L^2}{\partial \ln \tau_1^{\Delta\nu}(L_0)/\partial L} \quad (12)$$

243

244 Moreover, one can notice that as $-1/k_{p,1} \cdot \partial \ln \tau_1^{\Delta\nu}(L)/\partial L$ is a Laplace transform, it is log-convex
 245 [18]. This provides, following the definition Eq. (12), that $s_1(L_0)$ is a positive and decreasing
 246 function of L_0 :

$$247 \quad 0 \leq s_1(L_0) \leq s_1(0) \quad (13)$$

249
 250 The derivation provided in this section shares many similarities with SNB models, as the two
 251 approaches are founded on the same starting point (a spectrum is assumed to be the superposition
 252 of N statistically independent overlapping spectral lines). However, as our objective here is to
 253 provide proofs of Eqs. (11) and (13), that are the only results needed for our development, all
 254 additional assumptions required by SNB models [1] (choice of a distribution of linestrengths and of
 255 a line profile, assumption that spectral lines are fully included in $\Delta\nu$ and that the number of spectral
 256 line N tends towards infinity) are not needed. It can be checked easily that the property “
 257 $-1/k_{p,1} \cdot \partial \ln \tau_1^{\Delta\nu}(L)/\partial L$ is a Laplace transform” applies to standard SNB model formulations, as
 258 particular cases of the present derivation.

259

260 **2.2. The concept of quasi-scaled spectra**

261 This section is devoted to a description of the concept of quasi-scaled spectra, used in the present
 262 work to construct an approximation of $\ell_1 \circ \tau_2^{\Delta\nu}(L)$ as described in section 2.3.

263 The concept of quasi-scaled spectra is among the most important ones in the ℓ -distribution approach,
 264 at least as essential as that of “correlated” spectra in non-uniform k -distribution modeling, as they
 265 both consist of statistical models of relationship between gas spectra in distinct thermophysical
 266 states. We say that two spectra represented respectively by an index 1 and 2 are quasi-scaled if, by
 267 writing $\kappa_{\nu,2} = \kappa_{\nu,1} \cdot u_\nu$ where u_ν is the non-constant scaling coefficient between the two gas spectra,

268 $\kappa_{v,1}$ and u_v are statistically independent. In this case, we obtain the following relationship between
 269 the transmissivities in the two states averaged over the spectral interval $\Delta\nu$:

270

$$271 \quad \tau_2^{\Delta\nu}(L) = \frac{1}{\Delta\nu} \int_{\Delta\nu} \exp(-\kappa_{v,2}L) d\nu = \int_0^1 \tau_1^{\Delta\nu}[u(\xi)L] d\xi \quad (14)$$

272

273 where $u(\xi)$, $\xi \in [0,1]$, is the inverse of the distribution function $F(u)$ of the spectral variable u_v .

274 The concept of quasi-scaling was studied in depth in Ref. [19]. In this reference, it was verified that

275 within the assumption of quasi-scaled spectra, an approximation of the transmissivity of a non-

276 uniform path involving both states 1 and 2 can be obtained as:

277

$$278 \quad \tau_{12}^{\Delta\nu}(L_1, L_2) = \int_0^1 \tau_1^{\Delta\nu}[L_1 + u(\xi)L_2] d\xi \approx \tau_1^{\Delta\nu}[L_1 + \ell_1 \circ \tau_2^{\Delta\nu}(L_2)] \quad (15)$$

279

280 where ℓ_1 is the inverse of the transmissivity function $\tau_1^{\Delta\nu}$ defined as $\ell_1 \circ \tau_1^{\Delta\nu}(L) = L$ for any gas path

281 length L . This formalism is almost the same as a technique proposed by Godson in the 50s [20] but,

282 due to the quasi-scaling assumption, the present approach is more restrictive as explained in Refs.

283 [16,17]. A theoretical justification of the approximation set by Eq. (15) is provided later in this

284 paper.

285

286 **2.3. Lévy-Khintchine representation of $\ell_1 \circ \tau_2(L)$ in the case of weakly non-constant scaling**

287 **coefficients $u(\xi) \in [u_{\min}, u_{\max}]$, $\xi \in [0,1]$ where $|u_{\max} - u_{\min}| \ll u_{\min}$.**

288 Several analytical functional forms for $\ell_1 \circ \tau_2^{\Delta\nu}(L)$ were studied recently [9,19]. They were however

289 restricted to rather simple cases (gas spectrum in state 1 was gray, for instance, in Ref. [19]). In the

290 next two sections we consider the general situation. For this purpose, we first treat an intermediate

291 problem for which the non-constant scaling coefficient $u(\xi)$, $\xi \in [0,1]$ takes values inside a small
 292 interval $[u_{\min}, u_{\max}]$, $|u_{\max} - u_{\min}| \ll u_{\min}$. The corresponding result is then generalized to arbitrary
 293 situations.

294
 295 We consider a value ξ inside the interval $[0,1]$. The transmissivity $\tau_1^{\Delta v} [u(\xi)L]$ can be written as:

$$296 \quad \tau_1^{\Delta v} [u(\xi)L] = \exp \left[\int_0^{u(\xi)L} \frac{\partial \ln \tau_1^{\Delta v}(L')}{\partial L'} dL' \right] = \exp \left[\int_0^{u_{\min}L} \frac{\partial \ln \tau_1^{\Delta v}(L')}{\partial L'} dL' \right] \cdot \exp \left[\int_{u_{\min}L}^{u(\xi)L} \frac{\partial \ln \tau_1^{\Delta v}(L')}{\partial L'} dL' \right] \quad (16)$$

297
 298
 299
 300 or equivalently:

$$301 \quad \tau_1^{\Delta v} [u(\xi)L] = \tau_1^{\Delta v}(u_{\min}L) \cdot \exp \left[\int_{u_{\min}L}^{u(\xi)L} \frac{\partial \ln \tau_1^{\Delta v}(L')}{\partial L'} dL' \right] \quad (17)$$

302
 303 The change of variable $v(\xi) = u(\xi) - u_{\min}$ (notice from this definition that $v(\xi) \geq 0$) together with
 304 Eq. (11) then provides (function s_1 is defined in Section 2.2):

$$305 \quad \tau_1^{\Delta v} [u(\xi)L] \approx \tau_1^{\Delta v}(u_{\min}L) \cdot \exp \left(\frac{\partial \ln \tau_1^{\Delta v}(u_{\min}L)}{\partial L'} \cdot \frac{1 - \exp[-s_1(u_{\min}L) \cdot v(\xi)L]}{s_1(u_{\min}L)} \right) \quad (18)$$

306
 307
 308 As the width of the interval $[u_{\min}, u_{\max}]$ is assumed small, one can write:

$$309 \quad \tau_1^{\Delta v}(u_{\min}L) \geq \tau_1^{\Delta v} [u(\xi)L] \geq \tau_1^{\Delta v}(u_{\max}L) \approx \tau_1^{\Delta v}(u_{\min}L) \quad (19)$$

310
 311

312 that allows using the following approximation:

$$313 \quad \ln \left(\frac{\tau_1^{\Delta\nu} [u(\xi)L]}{\tau_1^{\Delta\nu} (u_{\min}L)} \right) \approx -1 + \frac{\tau_1^{\Delta\nu} [u(\xi)L]}{\tau_1^{\Delta\nu} (u_{\min}L)} \quad (20)$$

314

315 Eq. (20) provides, when combined with Eq. (18):

316

$$317 \quad \tau_1^{\Delta\nu} [u(\xi)L] \approx \tau_1^{\Delta\nu} (u_{\min}L) \cdot \left(1 + \frac{\partial \ln \tau_1^{\Delta\nu} (u_{\min}L)}{\partial L'} \frac{1 - \exp[-s_1(u_{\min}L) \cdot v(\xi)L]}{s_1(u_{\min}L)} \right) \quad (21)$$

318

319 This estimate of the transmissivity $\tau_1^{\Delta\nu} [u(\xi)L]$ can then be plugged into Eq. (14) to yield:

320

$$321 \quad \tau_2^{\Delta\nu} (L) \approx \tau_1^{\Delta\nu} (u_{\min}L) \cdot \left(1 + \frac{\partial \ln \tau_1^{\Delta\nu} (u_{\min}L)}{\partial L'} \cdot \int_0^1 \frac{1 - \exp[-s_1(u_{\min}L) \cdot v(\xi)L]}{s_1(u_{\min}L)} d\xi \right) \quad (22)$$

322

323 This result can be equivalently written in terms of the inverse of the band averaged transmissivity

324 $\tau_1^{\Delta\nu}$ as:

$$325 \quad \ell_1 \circ \tau_2^{\Delta\nu} (L) \approx u_{\min}L + \int_0^1 \frac{1 - \exp[-s_1(u_{\min}L) \cdot v(\xi)L]}{s_1(u_{\min}L)} d\xi \quad (23)$$

326

327 Using Eq. (23) directly is not convenient because it involves the quantity $s_1(u_{\min}L)$ that depends on

328 the gas path length L . No simple model for $s_1(u_{\min}L)$ seems to be available in the general frame

329 (see its definition, Eq. (12)). Consequently, in order to simplify the model, we rewrite Eq. (23) as:

330

$$331 \quad \ell_1 \circ \tau_2^{\Delta\nu} (L) \approx u_{\min}L \cdot [1 + \varphi(L)] + \int_0^1 \frac{1 - \exp[-s_1(0) \cdot v(\xi)L]}{s_1(0)} d\xi \quad (24)$$

332 where:

$$333 \quad \varphi(L) = \frac{1}{u_{\min} L} \cdot \left\{ \int_0^1 \frac{1 - \exp[-s_1(u_{\min} L) \cdot v(\xi) L]}{s_1(u_{\min} L)} d\xi - \int_0^1 \frac{1 - \exp[-s_1(0) \cdot v(\xi) L]}{s_1(0)} d\xi \right\} \quad (25)$$

334

335 The difference between the two integrals inside the brackets in Eq. (25) can be written as a double
336 integral to yield:

337

$$338 \quad 0 \leq \varphi(L) = \frac{1}{u_{\min}} \cdot \int_0^1 v(\xi) \cdot \underbrace{\left[\int_0^1 (\exp[-s_1(u_{\min} L) \cdot v(\xi) Lt] - \exp[-s_1(0) \cdot v(\xi) Lt]) dt \right]}_{\leq 1} d\xi \leq \frac{\bar{u} - u_{\min}}{u_{\min}} \ll 1$$

$$339 \quad (26)$$

340

341 The first inequality (on the left, viz. positivity) arises directly from application of the relationship
342 Eq. (13) that provides, for any (obviously positive value) of the product $v(\xi)Lt$,

343 $\exp[-s_1(L_0 = u_{\min} L) \cdot v(\xi) Lt] \geq \exp[-s_1(0) \cdot v(\xi) Lt]$. In Eq. (26), $\bar{u} = \int_0^1 u(\xi) d\xi$ is the mean

344 scaling coefficient equal to the ratio of the Planck mean absorption coefficients in states 1 and 2
345 (value at the optically thin limit).

346

347 Eq. (26) proves that Eq. (24) can be reasonably approximated as:

348

$$349 \quad \ell_1 \circ \tau_2^{\Delta\nu}(L) \approx u_{\min} L + \int_0^1 \frac{1 - \exp[-s_1(0) \cdot v(\xi) L]}{s_1(0)} d\xi \quad (27)$$

350

351 that does not depend anymore on $s_1(u_{\min} L)$. The parameter $s_1(0)$ that appears in Eq. (27) can be

352 estimated directly from LBL data following its definition:

353

$$354 \quad s_1(0) = -\frac{\partial^2 \ln \tau_1^{\Delta\nu}(0)/\partial L^2}{\partial \ln \tau_1^{\Delta\nu}(0)/\partial L} = \frac{k_{SP,1}}{k_{P,1}} - k_{P,1}, \quad k_{P,1} = \frac{1}{\Delta\nu} \cdot \int_{\Delta\nu} \kappa_{\nu,1} d\nu, \quad k_{SP,1} = \frac{1}{\Delta\nu} \cdot \int_{\Delta\nu} (\kappa_{\nu,1})^2 d\nu$$

355 (28)

356

357 In Eq. (28), k_p is the Planck mean absorption coefficient, whereas k_{SP} denotes Super Planck mean
 358 absorption coefficient [1]. These two mean absorption coefficients are both calculated in state 1 in
 359 Eq. (28).

360

361 One can notice here that if the same assumptions and steps as used in this section are used to evaluate
 362 $\ell_1 \circ \tau_{12}(L_1, L_2)$, where $\tau_{12}(L_1, L_2)$ is defined as the first term at the RHS in Eq. (15), one obtains:

363

$$364 \quad \ell_1 \circ \tau_{12}^{\Delta\nu}(L_1, L_2) \approx L_1 + u_{\min} L_2 + \int_0^1 \frac{1 - \exp[-s_1(0) \cdot \nu(\xi) L_2]}{s_1(0)} d\xi = L_1 + \ell_1 \circ \tau_2^{\Delta\nu}(L_2) \quad (29)$$

365

366 Proving this statement mostly requires changing all the occurrences $u_{\min} L$ (resp. $u(\xi) L$) by
 367 $L_1 + u_{\min} L_2$ (resp. $L_1 + u(\xi) L_2$) in Eqs. (16-26). This modification only affects the definition of
 368 function φ . This function remains however small compared to 1, yielding Eq. (29). This provides,
 369 *a posteriori*, a theoretical justification of the approximation set by the second part of Eq. (15).

370

371 Eq. (27) also shows that $\ell_1 \circ \tau_2^{\Delta\nu}(L)$ is a Bernstein function [21] (also called Laplace exponent of a
 372 Lévy subordinator in statistical science), viz. a function whose derivative is a Laplace transform.
 373 Formula (27) is the Lévy-Khintchine representation of this Bernstein function. It has the same
 374 mathematical form in the general case as in the previously treated simpler situations of Refs. [9,19].

375 This appartenance of $\ell_1 \circ \tau_2^{\Delta\nu}(L)$ to the family of Bernstein functions is central in the extension of
 376 the case treated in this section, that implicitly assumes that the two states 1 and 2 are close to each
 377 other, to more general scenarios. This generalization is discussed in the next section.

378

379 **2.4. General case**

380 It was shown in Ref. [19] that the quasi-scaling assumption is reasonable when the thermophysical
 381 states 1 and 2 are close to each other. This constraint is not verified in general radiative transfer
 382 calculations for which large gradients of temperature, pressure and species concentrations may be
 383 encountered along a radiation path. However, for any given couple of states 1 and 2, we can assume
 384 that there exists a sequence of N (finite) close intermediate states $\{1\}=1, \{2\}, \dots, \{N-1\}, \{N\}=2$ such
 385 that gas spectra between any couple of states $\{i\}$ and $\{i+1\}$ follow the assumptions of section 2.3.
 386 In this case, using these intermediate states, one can write:

387

$$388 \quad \ell_1 \circ \tau_2^{\Delta\nu}(L) = \ell_{\{1\}} \circ \tau_{\{N\}}^{\Delta\nu}(L) = \ell_{\{1\}} \circ \tau_{\{2\}}^{\Delta\nu} \left[\ell_{\{2\}} \circ \tau_{\{3\}}^{\Delta\nu} \left(\dots \left[\ell_{\{N-1\}} \circ \tau_{\{N\}}^{\Delta\nu}(L) \right] \right) \right] \quad (30)$$

389

390 *i.e.*, function $\ell_1 \circ \tau_2^{\Delta\nu}(L)$ can be constructed as the composition of $N-1$ intermediate functions. As
 391 gas spectra between any couple of states $\{i\}$ and $\{i+1\}$ follow the assumptions of section 2.3, by
 392 construction, all functions $\ell_{\{i\}} \circ \tau_{\{i+1\}}^{\Delta\nu}$, $i=1, \dots, N-1$ in Eq. (30) are of the Bernstein type. Their
 393 combination is thus also a Bernstein function [21], viz. of the form of Eq. (27). This property on
 394 combinations of Bernstein functions thus allows generalizing very straightforwardly our physical
 395 model of section 2.3 to more general situations. Notice that if the functional form of the solution is
 396 known at this level, the set of states $\{2\}, \dots, \{N-1\}$ is unknown. Its specification can be avoided by
 397 using tools from statistical learning, as explained in the next section. The method proposed to derive
 398 the model parameters from high resolution LBL data is adapted from Ref. [9] and is first described
 399 in this specific application context. The model set by Eqs. (15,27) is then applied to atmospheric

400 calculations. Comparisons with reference LBL calculations over non-uniform atmospheric paths are
401 then provided.

402

403

3. APPLICATION

404
405
406 In this section, we compare non-uniform atmospheric transmissivities obtained by application of
407 Eq. (27), combined with the propagative scheme of Eq. (3), with simulations obtained by the ℓ -
408 distribution method based on mapping functions, as used for instance in Ref. [17]. These two
409 methods are assessed against reference LBL calculations. Results of CKD models for various
410 numbers of gray gases are also provided for completeness. All model parameters are constructed
411 with the help of the same high resolution LBL data. Water vapor was chosen for the present analysis
412 and the Mid Latitude Summer profile was selected. This profile is discretized in 49 uniform layers
413 between 0 km (ground) and 120 km (top of the atmosphere). The METimage / EPS-SG channel
414 VII-20 optical filter (that defines a spectral range that extends from $10579.7715 \text{ cm}^{-1}$ up to
415 $11312.2172 \text{ cm}^{-1}$) is considered in the calculations. The spectral resolution of the LBL calculation
416 is 0.01 cm^{-1} (the spectral band thus contains 73,245 values of wavenumbers). Section 3.1 is devoted
417 to the analysis of two-layers systems. Its main objective is to demonstrate the relevance of Eq. (27)
418 for the calculation of $\ell_1 \circ \tau_2^{\Delta\nu}(L)$ and to show how the unknown quantities in the model can be
419 obtained by regression using a standard algorithm from the machine learning community. Section
420 3.2 describes preliminary results of a prototype code used to extend the results of section 3.1. to real
421 non-uniform atmospheres made of n higher than 2 uniform sub-layers. In the same section 3.2,
422 additional test cases related to oxygen in the A-band as well as a mixture of CO_2 , H_2O and O_3 are
423 also treated. For these supplementary cases, the optical filters of METimage / EPS-SG channel VII-
424 16 and METimage / EPS-SG channel VII-39 were selected respectively for O_2 and the CO_2 - H_2O -
425 O_3 mixture. The test cases considered are limited to purely absorbing atmospheres but extension of
426 the method to scattering atmospheres can be made in exactly the same way as described in Ref. [17].

427

428

429

430 **3.1. Analysis of two-layers configurations**

431 The objective of the present section is to study the validity of the functional form set by Eq. (27) to
 432 treat two-layers configurations. For this purpose, the estimation of the unknown coefficients of the
 433 model in the case of two-layers systems, that is to say the parameters u_{\min} and $v^*(\xi)$, is first made
 434 by applying the following steps:

435 Step 1/ for regression purpose, we select the following approximation of $\ell_1 \circ \tau_2^{\Delta v}(L)$ derived in
 436 Appendix II (Eq. (II.4)) and rigorously equivalent to Eq. (27):

437

$$438 \quad \ell_1 \circ \tau_2^{\Delta v}(L) \approx u_{\min} L + (\bar{u} - u_{\min}) \cdot \int_0^1 \frac{1 - \exp[-s_1(0) \cdot v^*(\xi)L]}{s_1(0) \cdot v^*(\xi)} d\xi \quad (31)$$

439

440 that we approximate, using a Gauss Legendre quadrature to calculate the integral that appears at the
 441 RHS (see Eq. (II.5)), as:

442

$$443 \quad \ell_1 \circ \tau_2^{\Delta v}(L) \approx \ell_1 \circ \tau_2^{\Delta v}[u_{\min}, v^*(x_i); L] = u_{\min} L + (\bar{u} - u_{\min}) \cdot \sum_{i=1}^N \omega_i \frac{1 - \exp[-s(0) \cdot v^*(x_i)L]}{s(0) \cdot v^*(x_i)} \quad (32)$$

444

445 The order of the Gauss Legendre quadrature (whose weights and nodes are written ω_i and x_i
 446 respectively) was set to 16, as suggested in Ref. [9]. The main interest of this choice of functional
 447 form for $\ell_1 \circ \tau_2^{\Delta v}(L)$ is that the initialization of the parameters to adjust is straightforward, as
 448 discussed in Appendix II.

449 Step 2/ a loss function \mathcal{L} is defined:

450

451
$$\mathcal{L}[u_{\min}, v^*(x_i)] = \int_0^{+\infty} \left[\tau_2^{\Delta v}(L) - \tau_1^{\Delta v}(\ell_1 \circ \tau_2^{\Delta v}[u_{\min}, v^*(x_i); L]) \right]^2 d\tau_2^{\Delta v}(L) \quad (33)$$

452
$$= \int_0^1 \left[\xi - \tau_1^{\Delta v}(\ell_1 \circ \tau_2^{\Delta v}[u_{\min}, v^*(x_i); \ell_2(\xi)]) \right]^2 d\xi$$

453 The integral from 0 to 1 is calculated, using a uniform discretization of the unit interval, as:

454

455
$$\mathcal{L}[u_{\min}, v^*(x_i)] = \frac{1}{J} \sum_{j=1}^J \left[\xi_j - \tau_1^{\Delta v}(\ell_1 \circ \tau_2^{\Delta v}[u_{\min}, v^*(x_i); \ell_2(\xi_j)]) \right]^2, \quad \xi_j = \frac{j}{J} \quad (34)$$

456

457 Application of the previous definition of the loss function requires evaluating the set of lengths

458 $\ell_2(\xi_j) = \ell_2(j/J)$, $j=1, J$. These quantities are calculated using high resolution (200,000 points)

459 mapping functions together with the “standard” ℓ -distribution model formulation (see Ref. [15] for

460 details). Mapping functions at the same high resolution are also used to calculate the transmissivity

461 functions in state 1, *i.e.*, $\tau_1^{\Delta v}$, as required by Eq. (34). In our calculations, the total number of lengths

462 along the transmission curves was set to $J=20,000$.

463 Step 3/ the loss function is minimized using ADAM’s [22] method. ADAM is a first-order gradient

464 optimization method based on adaptative estimates of lower-order moments. An ADAM code using

465 explicit gradients (simple to evaluate analytically using the definitions of Eqs. (32,34)) was written

466 in Fortran 90. The learning rate was set to 0.5 (rather standard value used in ADAM) and a simple

467 learning decay rate strategy was applied (for every step along the learning process where the loss

468 function increases, the learning rate is decreased by 0.05 percent – this simple strategy was found

469 to accelerate the convergence of the optimization process). The solution is attained when the loss

470 function between two successive iterations does not change from more than 10^{-4} %. In general, this

471 level of convergence is attained after a few dozen of thousands of iterations. Notice that in this

472 section, our objective is to validate Eq. (27), reason why such a strong convergence criterion was

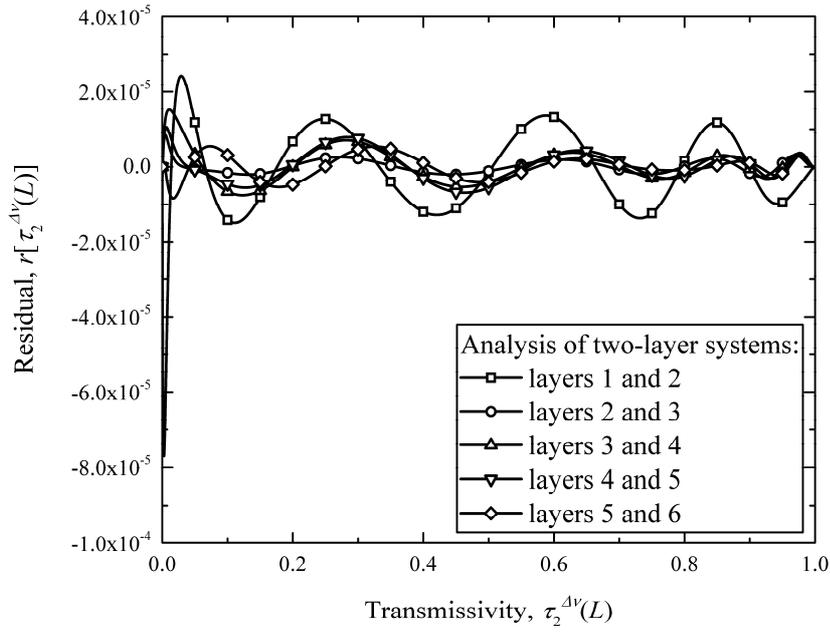
473 selected. In the optimization code, the following constraints are set to follow the definition of this
474 variable: $v^*(x_i) > 0, i = 1, N$.

475

476 Results of the model after optimization of its parameters are depicted in Figures 1, 2 and 3 for layers
477 between 0 (ground surface) and 16 km. Preliminary numerical tests (not depicted here) have shown
478 that, in these cases, errors of more than a few percent of the full transmission scale can be observed
479 if a single scaling coefficient (either based on the optically thin limit, viz. defined as the ratio of
480 Planck mean absorption coefficients, or at the optically thick limit, for which the minimum value of
481 scaling coefficient over the band is obtained) is used: gas spectra are thus not rigorously scaled in
482 the configurations treated. However, in all cases of Figure 1-3, the residual defined as the difference
483 $r[\tau_2^{\Delta\nu}(L)] = \tau_2^{\Delta\nu}(L) - \tau_1^{\Delta\nu}(\ell_1 \circ \tau_2^{\Delta\nu}[u_{\min}, v^*(x_i); L])$ is lower than 10^{-4} indicating the relevance of
484 the functional form set by Eq. (II.5), and thus as an extension of the model of Eq. (27), to calculate
485 $\ell_1 \circ \tau_2^{\Delta\nu}(L)$. The quality of the approximation shows, indirectly, that if all the assumptions made to
486 construct Eq. (27) are probably not rigorously verified, they are sufficiently realistic to provide
487 relevant and accurate estimates of $\ell_1 \circ \tau_2^{\Delta\nu}(L)$.

488

489

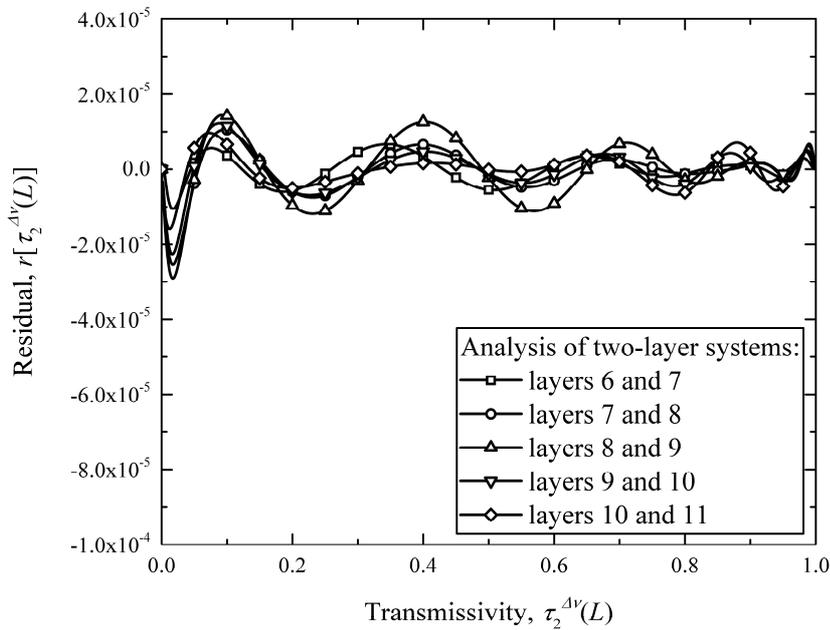


490

491 Figure 1. Residuals at the end of the fitting process for two-layers systems (for layers 1 to 5) – loss
 492 function given by Eq. (33).

493

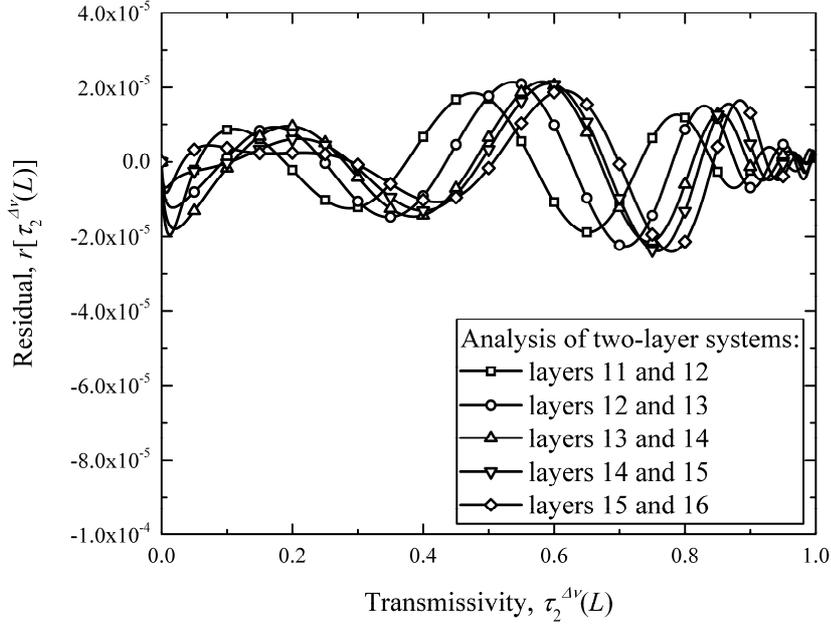
494



495

496 Figure 2. Residuals at the end of the fitting process for two-layers systems (for layers 6 to 10) –
 497 loss function given by Eq. (33).

498



499

500 Figure 3. Residuals at the end of the fitting process for two-layers systems (for layers 11 to 15) –
 501 loss function given by Eq. (33).

502

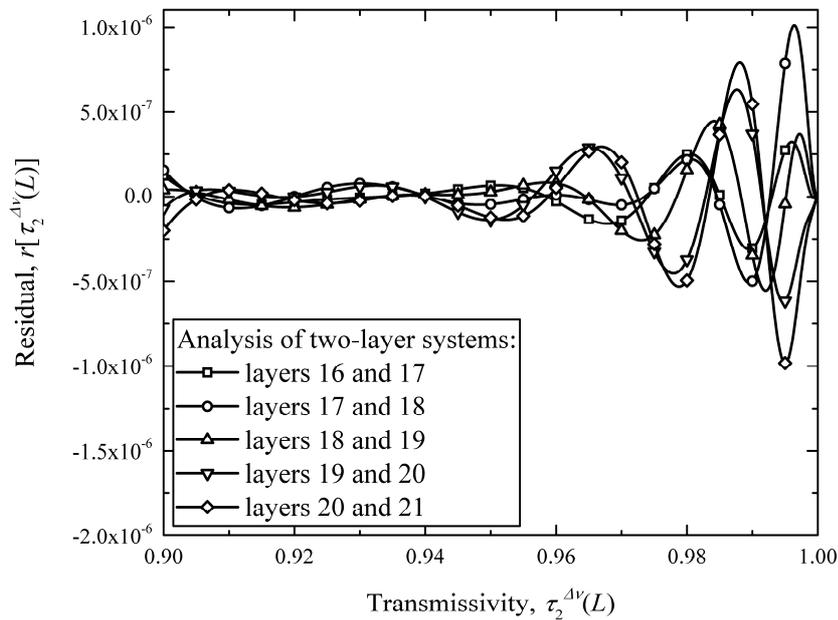
503 It is important to notice that at some locations along the non-uniform atmosphere, especially at high
 504 altitudes where the gaseous absorption is small, the set of input lengths $\ell_2(\xi_j) = \ell_2(j/J)$, $j=1, J$
 505 required for the optimization may take rather unrealistic values (for instance, for layer 20 which is
 506 20 km above the ground surface, the length required in the selected case of water vapor to absorb
 507 1% of the incident radiation, *i.e.*, the length to reach a value of transmissivity of 0.99, is higher than
 508 3000 km which is rather unrealistic for standard Earth remote sensing applications). In this case, the
 509 loss function is adapted to restrict the range of values of $\tau_2^{\Delta\nu}(L)$ to the Optically Thin Limit
 510 (exponent *OTL*):

511

$$512 \quad \mathcal{L}^{OTL}[u_{\min}, v^*(x_i)] = \int_{\tau_{\min}}^1 \left[\xi - \tau_1^{\Delta\nu} \left(\ell_1 \circ \tau_2^{\Delta\nu} [u_{\min}, v^*(x_i); \ell_2(\xi)] \right) \right]^2 d\xi \quad (35)$$

513

514 For the construction of Figure 4, that considers layers at altitudes located between 16 and 21 km,
 515 the value τ_{\min} was set to 0.9. This choice means that in the regression process, only the values of
 516 gas path lengths such that $\tau_2^{\Delta\nu}(L) \geq \tau_{\min} = 0.9$ are considered to evaluate the model's coefficients
 517 u_{\min} and $v^*(\xi)$. With this new loss function, very low residuals (absolute value of maximum is
 518 below 1.1×10^{-6}) are obtained at high altitudes as shown in Figure 4: the model set by Eq. (27) thus
 519 remains valid.



520
 521 Figure 4. Residuals at the end of the fitting process for two-layers systems (for layers 16 to 20) –
 522 Adjustment of the coefficients at the *OTL* (loss function defined by Eq. (35)).

523
 524 **3.2. Generalization to n -layer configurations**

525 In the previous section, we have restricted our analysis to two-layers systems. We have shown that
 526 the proposed functional form Eq. (27) provides, after an appropriate adjustment of the models'
 527 coefficients, an accurate approximation of $\ell_1 \circ \tau_2^{\Delta\nu}(L)$. In the present section, we treat the more
 528 general case of non-uniform atmospheres made of n (higher than 2) layers. For this purpose, we
 529 have combined the iterative scheme of Eq (3) with Eq. (27). This allows constructing a function of
 530 the n lengths encountered along a non-uniform atmosphere made of n distinct uniform layers. This

531 function can be trained over non-uniform path transmissivities following a logic similar to the one
532 described in the two-layers configurations of Section 3.1. But, in this case, the total number of
533 parameters to optimize is not 17 anymore but $17(n-1)$ (the value $n-1$ arises from the fact that in the
534 recurrent process, only $n-1$ functions of the form of Eq. (27) are used as shown in Eq. (3)). The
535 prototype code to generate the model's coefficients was written in Python using the PyTorch library.
536 The construction of the set of parameters for the n -layer system was done applying the following
537 steps (we provide here the main logic, as a proof of concept):

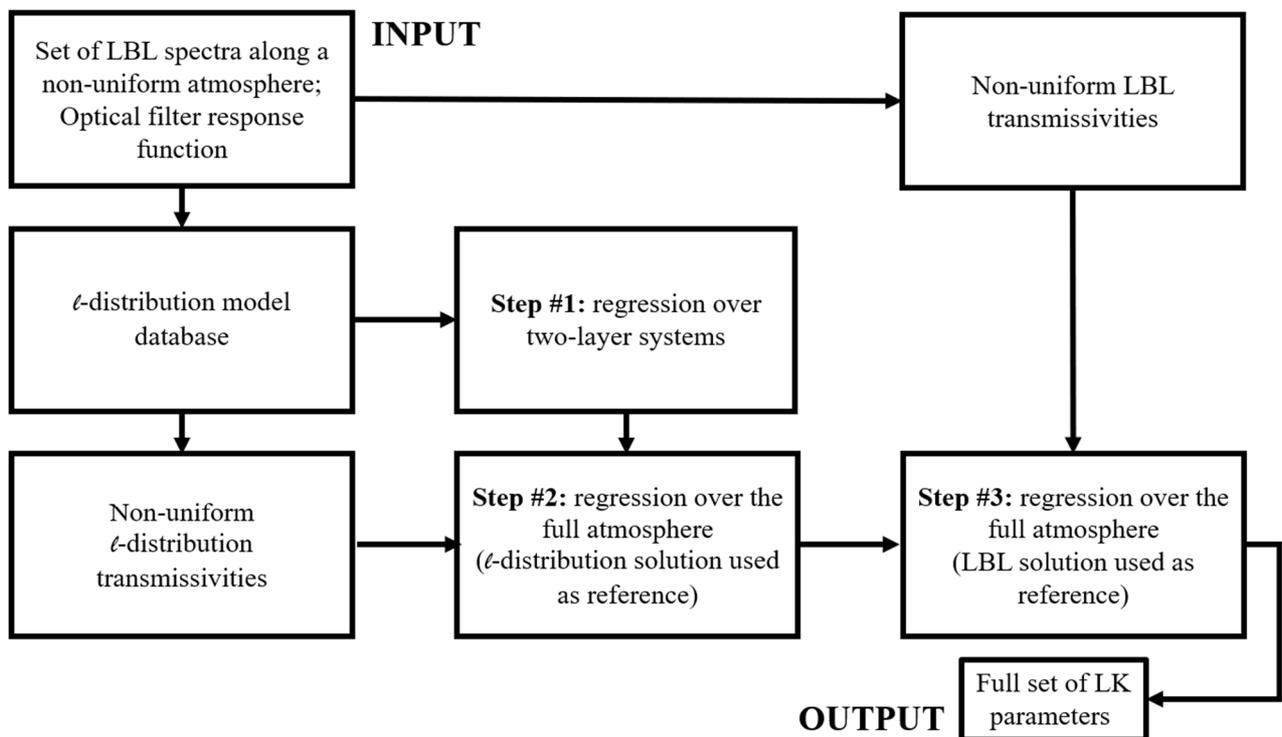
538 Step #1: a rough analysis of two-layers systems is first performed. ADAM method as provided in
539 PyTorch is used to improve the initial set (constructed in exactly the same way as in 3.1) but the full
540 convergence is not sought out in order to reduce the CPU cost. ADAM method is run for each couple
541 of layers for 500 iterations. This process is parallelized on 48 processors (one processor for each
542 couple of layers).

543 Step #2: at this level, we have a rough estimate of the model parameters. The full set of coefficients
544 is re-adjusted on non-uniform path transmissivities calculated with the help of the ℓ -distribution
545 method based on mapping functions. For this purpose, approximations of non-uniform path
546 transmissivities using the ℓ -distribution method are selected every 500 m (the full atmosphere has a
547 length of 120 km) and several relative air mass (1, 2, 4, 8, 16 and 24) are considered. The total
548 number of input data (values of non-uniform path transmissivities) used for the training process is
549 thus 1,440.

550 Step #3: non-uniform path transmissivities based on the ℓ -distribution method, as used in Step #2 to
551 train the full model, are replaced by transmissivities evaluated LBL. The training set is constructed
552 using the same method as for the ℓ -distribution training of Step #2 and also contains 1,440 values of
553 non-uniform path transmissivities. The main difference between Step #2 and Step #3 is that Step #2
554 uses non-uniform path transmissivities evaluated with the help of the ℓ -distribution model whereas
555 in Step #3 reference LBL calculations are used.

556 Step #4: once the set of parameters is considered sufficiently close to the optimal solution, the
 557 Bernstein functions Eq. (27) are used to update the mapping functions required by the ℓ -distribution
 558 method. This allows constructing a model that has the same accuracy as after Step #3 but with a
 559 minimal CPU cost. The model at this level is called Augmented ℓ -distribution. Additional details on
 560 the transition between Steps #3 and #4 are provided in Appendix III. A diagram that summarizes
 561 the various steps of the process is given in Figure 5.

562



563

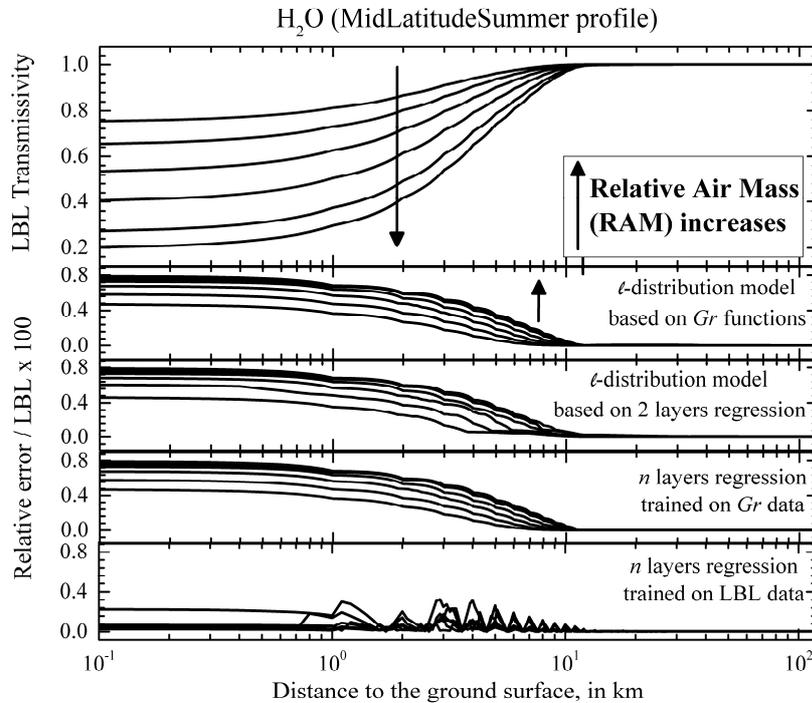
564

565 Figure 5. The several steps of the regression process. Input data (LBL spectra and optical filter)
 566 are on the top left – output data (coefficients of the Lévy-Khintchine (LK) formula all along the
 567 atmosphere) are on the bottom right. Step #4 is not depicted on the figure.

568

569 Figure 6 depicts results of comparisons for several values of Relative Air Mass, from 1 (single path
 570 in the non-uniform atmosphere) up to 24 (relationships between RAM values and angles between
 571 the incident radiation and the Earth surface can be found, for instance, in Ref. [17]). This figure

572 illustrates how the quality of the model (measured in terms of the relative error of the model
 573 compared with reference LBL calculations) evolves after each step of the process.
 574



575
 576 Figure 6. Evolution of the relative errors of the model at various stages of the training process.

577
 578 In Figure 6, the top layer shows LBL transmission curves over the non-uniform atmospheres. At a
 579 given distance from the ground surface (abscissa), it depicts the transmissivity of the path between
 580 the top of the atmosphere and this location / abscissa.

581 The second layer (starting again from the top) displays results of the “standard” t -distribution model
 582 based on mapping functions, as described in Refs. [15,17]. The accuracy of the method is rather
 583 satisfying, with relative differences lower than 0.8% when compared with reference LBL
 584 calculations.

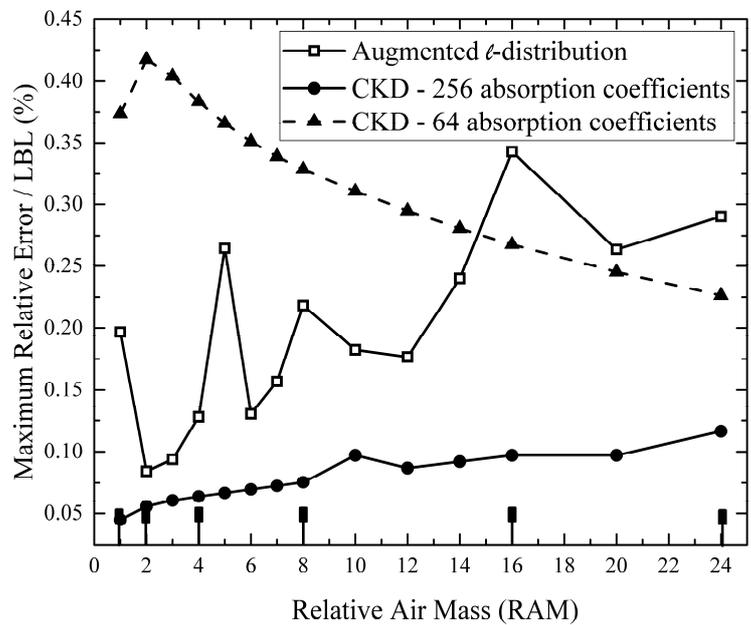
585 The third layer was constructed using directly the outputs of Step #1 (parameters adjustments on
 586 two-layers systems) combined with the iterative scheme of Eq. (3). Results are very close to those
 587 of the t -distribution model based on mapping functions (second layer), but not rigorously the same.

588 After Step #2 (fourth curve), the outputs of the “standard” ℓ -distribution model and the combination
589 of Eqs. (3,27) are virtually undistinguishable. This results can be explained by two factors: 1/ the
590 second training stage of Step #2 improves the initial learning stage of Step #1, which was probably
591 not pursued long enough to reach a full convergence, 2/ very likely, as only the set of lengths actually
592 used in the radiative transfer calculation are considered in the training, local adjustments are made
593 that allow improving further the model’s accuracy (in the same way, local adjustments at the *OTL*
594 were found to provide a slightly higher accuracy than if the full range of Eq. (33) was used in Section
595 3.1).

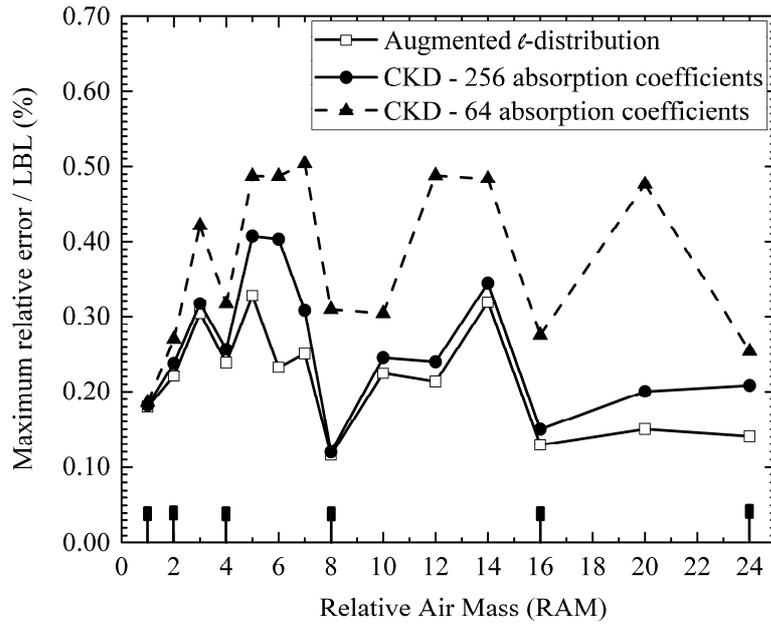
596 The fifth layer depicts results after Steps #3 and #4 (notice that when high resolution mapping
597 functions are constructed, as is the case here where 20,000 points are used for each updated mapping
598 function, results of Steps #3 and #4 are virtually identical). The gain in terms of accuracy compared
599 to the “standard” ℓ -distribution model is significant (a factor of about 2) but the two models used to
600 construct the second and fifth layers share the same CPU cost (after Step #4). This figure illustrates
601 the importance of the results described in the present work. Indeed, the use of the functional form
602 Eq. (27) derived on the present paper together with the iterative scheme of Eq. (3) allows
603 constructing a model of non-uniform path transmissivity that can reach a very high accuracy (of
604 about 0.35 % when compared with LBL data) at a very low CPU cost (the CPU time to generate a
605 full transmission curve, made of 1,200 values of non-uniform path transmissivities, is 0.1 ms on a
606 single core of an Intel Xeon Silver 4214R 2.4 GHz processor).

607
608 Eventually, one can notice that Figure 6 depicts results associated with the training set, obtained by
609 considering the values of relative air mass of 1, 2, 4, 8, 16 and 24. In order to evaluate how the
610 method extends to more general radiative transfer scenarios, we have generated model parameters
611 (using steps from #1 to #4) over 6 standard atmospheric profiles (Mid-Latitude Summer and Winter,
612 US standard, Tropical, Sub-Arctic Summer and Winter). Then, for a given value of relative air mass
613 between 1 and 24 by unit step, we have searched over the full transmission curve (as already said,

614 1,200 values of non-uniform path transmissivities are estimated for a given RAM) the maximum
 615 relative error of the model when compared with reference LBL calculations. For comparison
 616 purpose, we have made the same calculations for two CKD models based on Gauss-Legendre
 617 quadratures at orders 64 and 256. Results are depicted in Figure 7 for H₂O (same case as for Figure
 618 6), Figure 8 for O₂ and Figure 9 for the mixture (treated as a single gas). The set of RAMs used for
 619 the training is depicted by small vertical rectangles (connected to the RAM axis) in Figures 7-9.
 620



621
 622 Figure 7. Maximum relative errors / LBL for 6 standard atmospheres and values of relative air
 623 mass between 1 and 24 (case of water vapor).
 624
 625
 626

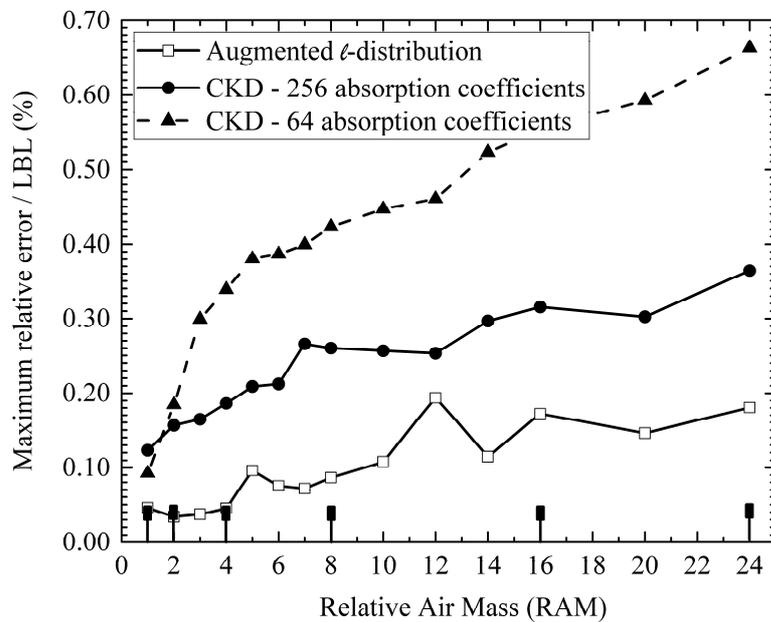


627

628 Figure 8. Maximum relative errors / LBL for 6 standard atmospheres and values of relative air
 629 mass between 1 and 24 (case of oxygen A-band).

630

631



632

633 Figure 9. Maximum relative errors / LBL for 6 standard atmospheres and values of relative air
 634 mass between 1 and 24 (case of CO₂-H₂O-O₃ mixture).

635

636 From Figures 7 to 9, several conclusions and comments can be drawn:

- 637 - The level of accuracy observed over the full set of RAMs indicates that the coefficients of
638 the sequence of Bernstein functions generalize well, even when applied to cases outside the
639 training set. The Augmented ℓ -distribution model provides an accuracy higher than 0.35%
640 for water vapor when assessed against reference LBL calculations, over the 6 standard
641 atmospheres. It is below 0.35% for O₂ and 0.20% for the mixture over the same set of cases.
- 642 - In most cases, for water vapor (Figure 7), the accuracy of the Augmented ℓ -distribution
643 method is lower than a CKD model with 256 gray gases but higher than a CKD model with
644 64 gray gases. This result cannot be considered as general (as illustrated by Figures 8 and 9)
645 because the CKD model performs particularly well for H₂O, as absorption by water vapor
646 only occurs in the first gaseous layers close to the ground where almost no gradient of
647 pressure and species concentrations is found (see the top layer of Figure 6: all the dynamics
648 of the transmission curves is below 10 km).
- 649 - Even if the Augmented ℓ -distribution model provides a slightly lower accuracy for H₂O than
650 the CKD model with 256 gray gases, it outperforms this method in terms of calculation cost.
651 Indeed, as shown in Ref. [17] where the “standard” ℓ -distribution method was compared with
652 CKD-256 k in realistic cloudy atmospheres, the CPU cost of Augmented ℓ -distribution
653 (which is the same as the ℓ -distribution model of Ref. [15,17] as only the mapping functions
654 differ between the two techniques) is about ten times lower than CKD-256 k . In many
655 applications, the 0.35% accuracy of Augmented ℓ -distribution method may be sufficient,
656 considering the significant gain in terms of CPU cost of this technique compared with the
657 more widely spread CKD model.
- 658 - For the oxygen and mixture cases (Figures 8 and 9), the Augmented ℓ -distribution model
659 provides a higher accuracy than both CKD-64 k and CKD-256 k methods at a lower
660 computational cost.

6. CONCLUSION

661
662
663 This paper is the continuation and generalization of a recent paper dedicated to the use of recurrent
664 methods to approximate non-uniform path transmissivities. A general formula was derived and
665 found to yield exactly the same recurrent structure as in our previous work. This formula involves
666 coefficients that have a clear physical meaning. However, they cannot be specified directly. A
667 training process using as input LBL transmission curves was thus developed to allow the
668 identification of the unknown coefficients of the proposed model. The non-uniform path
669 transmissivity approximation was then assessed against reference LBL calculations and found to
670 provide the same level of accuracy as the ℓ -distribution model with high-resolution (20,000 values)
671 mapping functions after a first training stage limited to the analysis of two-layers systems. A second
672 regression scheme was then added in which LBL data were used directly for the training of the
673 model's coefficients. After this second step, the model was found to provide very accurate
674 approximations of non-uniform path transmissivities, with errors lower than 0.35 % compared to
675 LBL in all the cases considered. Adding this second learning stage was shown to have no impact on
676 the CPU cost of the method, leading to an approximate technique that is among the most accurate
677 and fastest in the field.

678
679 Implications of the present results are numerous. From a practical perspective, the present technique
680 allows reducing significantly the memory cost of the ℓ -distribution method: indeed, look-up tables
681 of Gr functions using 20,000 points can be condensed into 17 parameters which can be useful in
682 terms of delivery of the model's parameters. Moreover, it also increases significantly its level of
683 accuracy (a factor of about 2 in terms of relative errors was obtained in this work). From a theoretical
684 perspective, the present developments show that the copula model used in the ℓ -distribution method
685 is of the Levy Subordinated type. The method developed here can be, a priori, extended to any other
686 LS-HAC copula models. The fact that a LS-HAC is obtained proves that the model embeds in its

687 structure some physical constraints about the derivatives of the transmissivity. Eventually, the
688 present work shows that the use of recurrent network structures to model non-uniform path
689 transmissivities is not limited to the simplified ℓ -distribution model treated in our previous paper but
690 that this property extends to LBL input data as considered here. Additionally, this result is not
691 limited to high temperature applications but it is shown to apply to atmospheric configurations too.

692

693

694

695

696 **ACKNOWLEDGEMENT**

697

698 This work was supported by the French National Research Agency under the grant ANR-20-CE04-
699 0005 ASGARD.

700

701

REFERENCES

702

703 [1] S. J. Young, *Band Model Theory of Radiation Transport*, The Aerospace Press, 2013.

704 [2] W. Malkmus, "Random Lorentz band model with exponential tailed S-1 line-intensity distribution
705 function", *J. Opt. Soc. Am.*, vol. 57, pp. 323-329, 1967.

706 [3] B.W. Webb, V.P. Solovjov, and F. André, "The spectral line based weighted-sum-of-gray-gases
707 (SLW) model for prediction of radiative transfer in molecucular gases", *Advances in Heat
708 Transfer*, vol. 51, pp. 207-298, 2019.

709 [4] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning: Data mining, Inference
710 and Prediction, Second Edition*, Springer Series in Statistics, 2017.

711 [5] A. del Aguila, D.S. Efremenko, T. Trautmann, "A review of dimensionality reduction techniques
712 for processing hyper-spectral optical signal", *Light and Engineering*, vol. 27, pp. 85-98, 2019.

713 [6] X. Liu, W.L. Smith, D.K. Zhou, A. Larar, "Principal component-based radiative transfer model for
714 hyperspectral sensors: theoretical concept", *Applied Optics*, vol. 45, pp. 201-208, 2006.

715 [7] F. André, C. Delage, C. Cornet, L. Croizé, Ph. Dubuisson, M. Galtier, "Computationally efficient
716 and accurate modelling of transmissivities of non-uniform paths through the mixture L-distribution
717 (MLD) approach", *in: Proceedings of the International Radiation Symposium, Thessaloniki, July
718 4-8, 2022.*

719 [8] P.G. Stegmann, B. Johnson, L. Moradi, B. Karpowicz, W. Mc Carty, "A deep learning approach
720 to fast radiative transfer", *J. Quant. Spectr. Rad. Transfer*, vol. 280, 108088, 2022.

- 721 [9] F. André, C. Cornet, C. Delage, Ph. Dubuisson, M. Galtier, “On the use of recurrent neural
722 networks for fast and accurate non-uniform gas radiation modeling”, *J. Quant. Spectr. Rad.*
723 *Transfer*, vol. 293, 108371, 2022.
- 724 [10] A. Royer, O. Farges, P. Boulet, D. Burot, “A new method for modeling radiative heat transfer
725 based on Bayesian artificial neural networks and Monte Carlo method in participating media”, *Int.*
726 *J. Heat Mass Transfer*, vol. 201, 123610, 2023.
- 727 [11] D.-A. Clevert, T. Unterthiner, S. Hochreiter, “Fast and accurate deep network learning by
728 exponential liner unit ‘ELUs)”, *in: Proceedings of the International Conference on Learning*
729 *Representation ICLR’16*, 2016.
- 730 [12] A. D. Jagtap, G. E. Karniadakis, “How important are activation functions in regression and
731 classification? A survey, performance comparison, and future directions”, *arXiv*, 2209.02681,
732 2022.
- 733 [13] J. Abbasi, P. O. Andersen, “Physical Activation Functions (PAFs): an approach for more efficient
734 induction of physics into Physics-Informed Neural Networks (PINNs)”, *arXiv*, 2205.14630, 2022.
- 735 [14] G. E. Karniadakis, I. G. Kevrekidis, L. Gu, P. Perdikaris, S. Wang, L. Yang, “Physics-informed
736 machine learning”, *Nature Review Physics*, vol. 3, pp. 422-440, 2021.
- 737 [15] F. André, “The ℓ -distribution method for modeling non-gray absorption in uniform and non-
738 uniform gaseous media”, *J. Quant. Spectr. Rad. Transfer*, vol. 179, pp. 19-32, 2016.
- 739 [16] F. André, “An analysis of the symmetry issue in the ℓ -distribution method of gas radiation in non-
740 uniform gaseous media”, *J. Quant. Spectr. Rad. Transfer*, vol. 190, pp. 78-87, 2017.
- 741 [17] F. André, C. Cornet, M. Galtier, Ph. Dubuisson, “Radiative transfer in the O₂ A-band – a fast and
742 accurate forward model based on the ℓ -distribution approach”, *J. Quant. Spectr. Rad. Transfer*, vol.
743 260, 107470, 2021.
- 744 [18] M. Merkle, “Completely Monotone Functions: A Digest”, *Analytic Number Theory,*
745 *Approximation Theory and Special Functions*, pp. 347-364, 2014.

- 746 [19] F. André, “Effective scaling factors in non-uniform gas radiation modeling”, *J. Quant. Spectr.*
747 *Rad. Transfer*, vol. 206, pp. 105-116, 2018.
- 748 [20] W.L. Godson, “The evaluation of infra-red radiative fluxes due to atmospheric water vapour”,
749 *Q. Journal Royal Meteorological Society*, vol. 79, pp. 367-379, 1953.
- 750 [21] R.L. Schilling, R. Song, Z. Vondraček, *Bernstein Functions: Theory and Applications*, Second
751 Edition, De Gruyter, 2012.
- 752 [22] D. P. Kingma, J. L. Ba, “Adam: a method for stochastic optimization”, *in: Proceedings of the 3rd*
753 *International Conference on Learning Representations, ICLR’15*, 2015.
- 754 [23] G. R. Schorack, *Probability for statisticians*, Springer text in statistics, Springer, 2000.
- 755 [24] C. Hering, M. Hofert, J.-F. Mai, M. Scherer, “Constructing hierarchical Archimedean copulas with
756 Levy subordinators”, *J. Multivariate Analysis*, vol. 101, pp. 1428-1433, 2010.
- 757 [25] R. B. Nelsen, *An introduction to copulas*, Second Edition, Springer series in statistics, Springer,
758 2006.
- 759 [26] H. Joe, *Multivariate models and dependence concepts*, Monograph on statistics and applied
760 probability 73, Springer Science+Business Media Dordrecht, 1997.
- 761 [27] M. Abramowitz, I. A. Stegun, *Handbook of Mathematical Functions*, Dover, 1965.
- 762 [28] F. André, R. Vaillon, “The k -moment method for modeling the blackbody weighted transmission
763 function for narrow and wide band radiative properties of gases”, *J. Quant. Spectr. Rad. Transfer*,
764 vol. 108, pp. 1-16, 2007.
- 765
- 766

APPENDIX I. Proof of Eq. (9)

767

768

769 The objective of this appendix is to provide a proof of Eq. (9). For this purpose, we first consider a
 770 single spectral line (index i) of the gas in state 1. The derivative of the transmissivity of this single line
 771 with respect to the gas path length L is (following a mathematical treatment similar to the one used in
 772 k -distribution methods):

773

$$774 \quad -\frac{\partial \tau_{1,i}^{\Delta\nu}}{\partial L} = \frac{1}{\Delta\nu} \cdot \int_{\Delta\nu} \kappa_{v,1,i} \exp(-\kappa_{v,1,i} L) d\nu = \int_0^{+\infty} \exp(-k^* \cdot L) \cdot \left[\frac{1}{\Delta\nu} \int_{\Delta\nu} \frac{\partial H(k^* - \kappa_{v,1,i})}{\partial k^*} \kappa_{v,1,i} d\nu \right] dk^* \quad (I.1)$$

775

776

777 where H is the Heaviside step function.

778 The sum over all spectral lines, *i.e.* over all indices i , then provides:

779

$$780 \quad \sum_{i=1}^N \left[-\frac{\partial \tau_{1,i}^{\Delta\nu}(L)}{\partial L} \right] = \int_0^{+\infty} \exp(-k^* \cdot L) \cdot \left\{ \sum_{i=1}^N \left[\frac{1}{\Delta\nu} \int_{\Delta\nu} \frac{\partial H(k^* - \kappa_{v,i})}{\partial k^*} \kappa_{v,i} d\nu \right] \right\} dk^* \quad (I.2)$$

781

782 Dividing both sides of Eq. (I.2) by the Planck mean absorption coefficient of the gas in state 1 yields:

783

$$784 \quad \frac{1}{k_{P,1}} \sum_{i=1}^N \left[-\frac{\partial \tau_{1,i}^{\Delta\nu}(L)}{\partial L} \right] = \int_0^{+\infty} \exp(-k^* \cdot L) \cdot \left\{ \frac{1}{k_{P,1}} \sum_{i=1}^N \left[\frac{1}{\Delta\nu} \int_{\Delta\nu} \frac{\partial H(k^* - \kappa_{v,i})}{\partial k^*} \kappa_{v,i} d\nu \right] \right\} dk^* \quad (I.3)$$

$$= \int_0^{+\infty} \exp(-k^* \cdot L) \cdot \frac{\partial}{\partial k^*} \left\{ \underbrace{\frac{1}{k_{P,1}} \sum_{i=1}^N \left[\frac{1}{\Delta\nu} \int_{\Delta\nu} H(k^* - \kappa_{v,i}) \kappa_{v,i} d\nu \right]}_{F(k^*)} \right\} dk^*$$

785

786 It can be readily checked that $F(k^*)$: 1/ is a strictly increasing function of the variables k^* , 2/ takes
 787 value 0 for $k^* = 0$ and 1 for $k^* = +\infty$. $F(k^*)$ is thus a distribution [23]. This allows rewriting (using
 788 the inverse $k^*(\xi)$ of $F(k^*)$ defined as $F[k^*(\xi)] = \xi$, $\xi \in [0,1]$):

789

$$790 \quad \frac{1}{k_{p,1}} \sum_{i=1}^N \left[-\frac{\partial \tau_{i,i}^{\Delta v}(L)}{\partial L} \right] = \int_0^1 \exp[-k^*(\xi) \cdot L] d\xi \quad (I.4)$$

791

792 Eq. (I.4) provides Eq. (9) when combined with Eq. (8).

793

794

795

APPENDIX II. Reformulations of Eq. (27) and their properties

796

797

798 The objective of this second appendix is to derive two equivalent formulations of Eq. (27). It is
 799 motivated by the observation that even if Eq. (27) has strong implications in terms of the statistical
 800 analysis of the model (indeed, as the combinations of successive inverses of transmissivities and
 801 transmissivities are Bernstein functions, application of the recurrent scheme of Eq. (3) is equivalent to
 802 the construction of a Lévy Subordinated Hierarchical Archimedean Copula LS-HAC [24], for which
 803 the so-called sufficient nesting condition [25,26] is naturally fulfilled - this “copula” view is not
 804 analyzed further in the present work but additional details can be found in Ref. [17]), Eq. (27) is neither
 805 the most convenient for regression purpose nor for a physical analysis of the model. Consequently, we
 806 derive here two equivalent formulations of Eq. (27) and illustrate how they can be used to provide an
 807 initialization of the regression scheme described in Section 3.1 and to a better understanding of the
 808 physical model described in this work.

809

810 For this purpose, we first modify Eq. (27) into:

811

$$812 \quad \iota_1 \circ \tau_2^{\Delta v} (L) \approx u_{\min} L + \int_0^1 \frac{1 - \exp[-s_1(0) \cdot v(\xi) L]}{s_1(0) \cdot v(\xi)} \cdot v(\xi) d\xi \quad (\text{II.1})$$

813

814 Eq. (II.1) can be equivalently rewritten as:

815

$$816 \quad \iota_1 \circ \tau_2^{\Delta v} (L) \approx u_{\min} L + \int_0^1 \frac{1 - \exp[-s_1(0) \cdot v(\xi) L]}{s_1(0) \cdot v(\xi)} d \left[\int_0^\xi v(\xi') d\xi' \right] \quad (\text{II.2})$$

817

818 that is to say, as $\int_0^1 v(\xi') d\xi' = \frac{1}{\Delta v} \cdot \int_{\Delta v} (u_v - u_{\min}) dv = \bar{u} - u_{\min}$:

819
$$\ell_1 \circ \tau_2^{\Delta v}(L) \approx u_{\min} L + (\bar{u} - u_{\min}) \cdot \int_0^1 \frac{1 - \exp[-s_1(0) \cdot v(\xi)L]}{s_1(0) \cdot v(\xi)} d \left[\underbrace{\frac{1}{\bar{u} - u_{\min}} \cdot \int_0^\xi v(\xi') d\xi'}_{F^*(\xi)} \right] \quad (\text{II.3})$$

820

821 The quantity $F^*(\xi)$ is a distribution over the unit interval. This provides:

822

823
$$\ell_1 \circ \tau_2^{\Delta v}(L) \approx u_{\min} L + (\bar{u} - u_{\min}) \cdot \int_0^1 \frac{1 - \exp[-s_1(0) \cdot v^*(\xi)L]}{s_1(0) \cdot v^*(\xi)} d\xi \quad (\text{II.4})$$

824

825 where $v^*(\xi) = v \circ (F^*)^{-1}(\xi)$, $\xi \in [0,1]$.

826 Using a Gauss Legendre quadrature at order N to approximate the integral over $[0,1]$ that appears

827 in Eq. (II.4), we obtain:

828

829
$$\ell_1 \circ \tau_2^{\Delta v}(L) \approx u_{\min} L + (\bar{u} - u_{\min}) \cdot \sum_{i=1}^N \omega_i \cdot \frac{1 - \exp[-s_1(0) \cdot v^*(x_i)L]}{s_1(0) \cdot v^*(x_i)} \quad (\text{II.5})$$

830

831 where ω_i , $i = 1, N$ and x_i , $i = 1, N$ are respectively the weights and nodes of the Gauss Legendre
832 quadrature.

833 Eq. (II.5) can be compared with Eqs. (19,29) from Ref. [9] that yields, in the case of gas spectra that
834 follow rigorously the assumptions of the SNB model for Lorentz lines with Malkmus' distribution of
835 linestrengths:

836

837
$$\ell_1 \circ \tau_2^{\Delta v}(L) \approx \frac{k_{P,2}\beta_2}{k_{P,1}\beta_1} L + \left(\bar{u} - \frac{k_{P,2}\beta_2}{k_{P,1}\beta_1} \right) \cdot \sum_{i=1}^N \omega_i \cdot \frac{1 - \exp[-u(x_i) \cdot 2\pi k_{P,2} / \beta_2 \cdot L]}{u(x_i) \cdot 2\pi k_{P,2} / \beta_2} \quad (\text{II.6})$$

838

839 where (see Ref. [9]) parameters $u(x_i)$, $i=1, N$ are obtained as solutions of the implicit equations (P
 840 is the incomplete gamma function [27]):

841

$$842 \quad P[1/2 ; u(x_i)] = x_i, \quad P(a ; u) = \frac{1}{\Gamma(a)} \cdot \int_0^u t^{a-1} \exp(-t) dt \quad (\text{II.7})$$

843

844 Based on this result, we can use the formal analogy between Eqs. (II.5) and (II.6) to initialize the
 845 regression scheme by setting, at the first step of the minimization process: $1/ u_{\min} \leftarrow \text{Min}(\kappa_{v,2}/\kappa_{v,1})$,

846 and $2/ v^*(x_i) \leftarrow \frac{\alpha \cdot \pi k_{P,2}}{\beta_2} \cdot \frac{1}{s_1(0)} \cdot u(x_i)$ with $\alpha = 2$ at the OTL (in this case the model of Eq. (II.6) is

847 almost exact [1,16,28]) and $\alpha = 4$ when the full range of lengths is used (this value of $\alpha = 4$ was
 848 found in practice to provide slightly more accurate results than $\alpha = 2$ when the loss function is defined

849 by Eq. (33)). Selecting the value $u_{\min} \leftarrow \frac{k_{P,2}\beta_2}{k_{P,1}\beta_1}$ as suggested by Eq. (II.6) was not found to yield, in

850 the treated cases, significant differences with $u_{\min} \leftarrow \text{Min}(\kappa_{v,2}/\kappa_{v,1})$. Additional details about the
 851 regression method are given in Section 3.1.

852

853 Moreover, Eq. (II.4) can be written as:

854

$$855 \quad \ell_1 \circ \tau_2^{\Delta v}(L) \approx u_{\min} L + (\bar{u} - u_{\min}) L \cdot \int_0^1 \frac{1 - \exp[-s_1(0) \cdot v^*(\xi) L]}{s_1(0) \cdot v^*(\xi) L} d\xi \quad (\text{II.8})$$

856

857 The quantity inside the integral at the RHS can be itself written as an integral to yield:

858

$$859 \quad \ell_1 \circ \tau_2^{\Delta v}(L) \approx u_{\min} L + (\bar{u} - u_{\min}) L \cdot \int_0^1 \int_0^1 \exp[-s_1(0) \cdot v^*(\xi) Lt] d\xi dt \quad (\text{II.9})$$

860 Introducing the path dependent / effective scaling coefficient $u(L)$ such that $\ell_1 \circ \tau_2^{\Delta\nu}(L) = u(L) \cdot L$,
 861 we eventually obtain (in the next equation as well as in Eq. (II.9), variable t is a dummy variable
 862 only required to reformulate the integral. It has no physical meaning and no unit):

863

$$864 \quad u(L) \approx u_{\min} + (\bar{u} - u_{\min}) \cdot \int_0^1 \int_0^1 \exp[-s_1(0) \cdot v^*(\xi) Lt] d\xi dt \quad (\text{II.10})$$

865

866 This relationship shows that the effective scaling coefficient is a decreasing function of the gas path
 867 length. It provides the proper asymptotic limits of the quasi-scaled model at the optically thin (\bar{u} for
 868 small values of L) and thick (u_{\min} for large values of L) limits. Moreover, as the double integral
 869 decreases with respect to L , we have $u(L) \leq \bar{u}$ and thus:

870

$$871 \quad \tau_2^{\Delta\nu}(L) = \tau_1^{\Delta\nu}[u(L) \cdot L] \geq \tau_1^{\Delta\nu}(\bar{u}L) \quad (\text{II.11})$$

872

873 This result is consistent with the following inequality (application of Jensen's inequality to the
 874 transmissivity function $\tau_1^{\Delta\nu}$):

875

$$876 \quad \tau_2^{\Delta\nu}(L) = \int_0^1 \tau_1^{\Delta\nu}[u(\xi) \cdot L] d\xi \geq \tau_1^{\Delta\nu}\left[\int_0^1 u(\xi) d\xi \cdot L\right] = \tau_1^{\Delta\nu}(\bar{u}L) \quad (\text{II.12})$$

877

878 **APPENDIX III. Analysis of the transition from Step #2 up to #4 of section 3.2**

879

880 The objective of this appendix is to provide details about the transition between the regression scheme
 881 of Step #2 described in section 3.2 (based on adjustment of the parameters on solution of a “standard”
 882 ℓ -distribution model), Step #3 (regression on non-uniform path LBL transmissivities) and Step #4
 883 (update of the mapping functions used in the “standard” ℓ -distribution formulation).

884

885 As can be seen on the fourth layer of Figure 6, the “standard” ℓ -distribution model provides an estimate
 886 of the non-uniform path transmissivity as:

887

$$888 \tau_{1..n}^{\Delta v}(L_1, \dots, L_n) \approx \tau_1^{\Delta v}(L_{1..n}) \quad (\text{III.1})$$

889

890 where $L_{1..n}$ is obtained by the following recurrent process:

891

$$892 \begin{cases} L_{nn} = L_n \\ L_{i..n} = L_i + \ell_i \circ \tau_{i+1}^{\Delta v}(L_{i+1..n}) \end{cases} \quad (\text{III.2})$$

893

894 in which (for instance for the couple of layers 1 and 2):

895

$$896 \ell_1 \circ \tau_2^{\Delta v}(L) \approx \ell_1 \circ \tau_2^{\Delta v}[u_{\min}, v^*(x_i); L] = u_{\min} L + (\bar{u} - u_{\min}) \cdot \sum_{i=1}^N \omega_i \frac{1 - \exp[-s(0) \cdot v^*(x_i) L]}{s(0) \cdot v^*(x_i)}$$

897 (III.3)

898

899 The third step of section 3.2 consists of an adjustment of all the parameters involved in the recurrent
 900 scheme on non-uniform LBL data. This produces a new estimate:

901

902

$$\tau_{1..n}^{\Delta v}(L_1, \dots, L_n) \approx \tau_1^{\Delta v}(L_{1..n}^+) \quad (\text{III.4})$$

903

904 where $L_{1..n}^+$ is obtained by the following recurrent process:

905

906

$$\begin{cases} L_{nn}^+ = L_n \\ L_{i..n}^+ = L_i + \lambda_i(L_{i+1..n}^+) \end{cases} \quad (\text{III.5})$$

907

908 in which (for instance for the couple of layers 1 and 2):

909

910

$$\lambda_1(L) \approx \lambda_1[u_{\min}^+, v^+(x_i); L] = u_{\min}^+ L + (\bar{u} - u_{\min}^+) \cdot \sum_{i=1}^N \omega_i \frac{1 - \exp[-s(0) \cdot v^+(x_i) L]}{s(0) \cdot v^+(x_i)} \quad (\text{III.6})$$

911

912 In general, as can be seen by comparing the fourth and fifth layers of Figure 6:

913

914

$$\tau_1^{\Delta v}(L_{1..n}) \neq \tau_1^{\Delta v}(L_{1..n}^+) \quad (\text{III.7})$$

915

916 Let us now consider the case for which all lengths are set to 0 except in layer i chosen arbitrarily in

917 $\{2, N\}$. We have:

918

919

$$L_{1..n} = L_{1..n}(L_1 = 0, L_2 = 0, \dots, L_i, \dots, L_n = 0) = [\ell_1 \circ \tau_2^{\Delta v}] \circ [\ell_2 \circ \tau_3^{\Delta v}] \circ \dots \circ [\ell_{i-1} \circ \tau_i^{\Delta v}](L_i) \quad (\text{III.8})$$

920

921 and thus:

922

923
$$\tau_1^{\Delta v}(L_{1..n}) = \tau_i^{\Delta v}(L_i) \quad (\text{III.9})$$

924

925 For the same set of lengths, we have:

926

927
$$L_{1..n}^+ = L_{1..n}^+(L_1=0, L_2=0, \dots, L_i, \dots, L_n=0) = \lambda_1 \circ \lambda_2 \circ \dots \circ \lambda_{i-1}(L_i) \quad (\text{III.10})$$

928

929 that leads to:

930

931
$$\tau_1^{\Delta v}(L_{1..n}^+) = \tau_1^{\Delta v} \circ \lambda_1 \circ \lambda_2 \circ \dots \circ \lambda_{i-1}(L_i) = \tau_i^+(L_i) \neq \tau_i^{\Delta v}(L_i) \quad (\text{III.11})$$

932

933 The use of the propagative scheme set by Eq. (III.5) instead of Eq. (III.2) does not allow recovering
 934 the transmission curves encountered along the non-uniform path but generates a new set of functions
 935 τ_i^+ of the gas path lengths L_i , $i=2, \dots, n$. These functions are obtained as the combination of
 936 Bernstein functions (that also provides a Bernstein function [21]) and of a Laplace transform (the
 937 true transmissivity in the first layer, as seen in Eq. (III.11)): τ_i^+ are consequently Laplace transforms
 938 too. These new functions are strictly decreasing and thus invertible. Let us write ℓ_i^+ these inverses.

939 From the definition of functions τ_i^+ we have:

940

941
$$\tau_i^+(L_i) = \tau_1^{\Delta v} \circ \lambda_1 \circ \lambda_2 \circ \dots \circ \lambda_{i-1}(L_i) = \tau_{i-1}^+ \circ \lambda_{i-1}(L_i) \quad (\text{III.12})$$

942

943 and thus:

944

945
$$\lambda_{i-1}(L_i) = \ell_{i-1}^+ \circ \tau_i^+(L_i) \quad (\text{III.13})$$

946 This shows that the recurrent scheme of Eq. (III.5) takes exactly the same form as the one given by
 947 Eq. (III.2). The same optimized code as used for the “standard” ℓ -distribution method can thus be
 948 used to apply the Augmented version, by simply replacing the look-up tables based on $\tau_i^{\Delta\nu}$, $i = 1..n$
 949 by look-up tables based on $\{\tau_1^{\Delta\nu}, \tau_2^+, \dots, \tau_n^+\}$. Calculation of these new look-up tables can be made
 950 by applying the definition of functions τ_i^+ given by Eq. (III.11).

951
 952 One can eventually check without difficulty that the corresponding model (based on the set
 953 $\{\tau_1^{\Delta\nu}, \tau_2^+, \dots, \tau_n^+\}$) follows the following constraints on the derivative of the non-uniform path
 954 transmissivity:

$$956 \quad -\frac{\partial \tau_1^{\Delta\nu}(L_{1..n}^+)}{\partial L_i} \leq k_{p,i} \cdot \tau_1^{\Delta\nu}(L_{1..n}^+), \quad i = 1, \dots, n \quad (\text{III.14})$$

957
 958 and:

$$960 \quad \frac{\partial^2 \tau_1^{\Delta\nu}(L_{1..n}^+)}{\partial L_i \partial L_j} \geq 0, \quad i = 1, \dots, n, \quad j = 1, \dots, n \quad (\text{III.15})$$

961
 962 Eq. (III.14), that follows here from the convexity property of Bernstein functions (see also Ref. [17]
 963 for additional details), has a long history in the development of band model theory. Indeed, in some
 964 configurations, the widely spread Curtis-Godson approximation was shown to fail to ensure this
 965 inequality. This led to the development of the Lindquist-Simmons approximation [1]. Eq. (III.15), that
 966 follows directly from the LS-HAC copula model, ensures the proper sign of the method when used for
 967 the calculation of net exchange rates, as discussed in Ref. [17].