The EMIT L3 Mass Abundance Calculation (Step 3): Physical Basis and Derivation of the Equations

Notes prepared for Z. Zhang
Based on the EMIT L3ASA ATBD V003 (Okin et al., JPL D-104480, July 2025)

June 2026

1 The problem Step 3 solves

Step 3 inherits from the L2B stage (Tetracorder; Clark et al., 2003) a set of detected minerals and their *band depths*: the fractional depth of each diagnostic absorption feature relative to the local continuum,

$$BD \;=\; 1 - \frac{R_b}{R_c}, \tag{1}$$

exactly as defined by Clark and Roush (1984), where R_b is the reflectance at the band center and R_c the interpolated continuum reflectance at the same wavelength.

Band depth is *not* proportional to mass fraction. In an intimate mixture, photons random-walk through a matrix of touching grains and sample many particles before escaping, so the apparent depth of a mineral's feature depends nonlinearly on its abundance, its grain size, and the optical properties of the spectrally bland host minerals (quartz and feldspar, which lack diagnostic VSWIR absorptions and are invisible to EMIT). This is the regime treated by Hapke (1981), whose bidirectional reflectance

$$r(\mu_0, \mu, g) \;=\; \frac{w}{4\pi}\,\frac{\mu_0}{\mu_0 + \mu}\Big[\big(1 + B(g)\big)\,p(g) + H(\mu_0)\,H(\mu) - 1\Big] \tag{2}$$

is controlled by the single-scattering albedo w; for an intimate mixture w is the cross-section-weighted average over components,

$$w \;=\; \frac{\sum_i M_i\, w_i/(\rho_i D_i)}{\sum_i M_i/(\rho_i D_i)}, \tag{3}$$

with M_i, ρ_i, D_i the mass fraction, density, and grain diameter of component i.

Rather than invert the full nonlinear model pixel by pixel, the EMIT L3 algorithm uses an equivalent but much simpler photon-path bookkeeping argument, following Clark et al. (2024), Section 11.4. The logic has five links, derived below.

2 Link 1: Band depth to mean optical path (Beer–Lambert)

Within the absorption feature, the only property distinguishing the band center from the continuum is the *extra* absorption of the diagnostic mineral. Writing the in-band reflectance as the continuum reflectance attenuated by Beer–Lambert extinction along the photon's cumulative path $\bar{d}_i$ inside grains of mineral i,

$$R_b = R_c\, e^{-k_i \bar{d}_i} \quad\Longrightarrow\quad t \equiv \frac{R_b}{R_c} = e^{-k_i \bar{d}_i}, \qquad BD = 1 - t, \tag{4}$$

and inverting,

$$\boxed{\bar{d}_i \;=\; \frac{\ln(1-BD)}{-\,k_i}} \tag{5}$$

Here k_i is the `delta_k` of ATBD Table 2: the *continuum-removed* absorption coefficient, $k(\lambda_{\text{band}}) - k(\lambda_{\text{continuum}})$, in cm^{-1}. Continuum removal is what makes the simple Beer–Lambert form defensible: scattering losses and broadband absorption are common to band and continuum and divide out in the ratio R_b/R_c.

A consistency check against optical constants: for hematite at 0.75 μm the imaginary index is $n_i \approx 0.04$, giving $\alpha = 4\pi n_i/\lambda \approx 6.7 \times 10^3$ cm^{-1}, in good agreement with the tabulated 6400 cm^{-1} for hematite-fine.

The implicit approximation is that absorption perturbs the photon path only weakly, so the path-length distribution is fixed by scattering geometry alone (the continuum random walk) and absorption merely attenuates along it. This is the weak-absorption limit of Hapke theory, appropriate for the small band depths typical of natural soils.

3 Link 2: Path to number of grain encounters

The photon accumulates the path $\bar{d}_i$ in discrete increments, one grain transit at a time. With D_i the grain diameter and an effective chord of $2D_i$ per encounter (a convention from Clark et al. 2024 that folds in internal reflections lengthening the in-grain path), the mean number of scattering events on mineral i is

$$\boxed{m_i \;=\; \frac{\bar{d}_i}{2D_i}} \tag{6}$$

4 Link 3: Encounters to penetration depth (diffusion)

Between absorption events the photon executes a nearly isotropic random walk with step length of order the grain diameter, so the depth probed below the surface grows as the square root of the number of steps, as in any diffusion process:

$$\boxed{T_i \;=\; \frac{\pi}{4}\, D_i \left(\frac{m_i}{2}\right)^{1/2}} \tag{7}$$

The numerical factors are order-unity geometric constants: $\pi D/4$ is the mean chord of a circular cross-section of diameter D (the average projected step), and the factor $1/2$ inside the root accounts for projection of the isotropic walk onto the vertical. For comparison, exact isotropy gives a variance factor of $1/3$ per step, i.e. $T \approx 0.58\, D\sqrt{m}$ versus $0.55\, D\sqrt{m}$ here, the same physics within the model's accuracy. T_i is the spectroscopic sampling depth: the thickness of the surface layer the measurement actually interrogates.

5 Link 4: Penetration depth to areal mass

The mass per unit surface area of a solid slab of thickness t and density ρ is simply $\rho\, t$. Equation (8) below applies this to an equivalent solid column of mineral i of thickness $3T_i$. The factor 3 arises because T_i is the RMS of a zero-mean vertical displacement (Eq. (7)), not the extent of the region actually sampled: the depth distribution of the photon ensemble has a tail, and the layer containing

essentially all of the sampled material extends to a few standard deviations. The 3 is the standard containment convention (three sigma captures 99.7% of a Gaussian):

$$\boxed{M_i \;=\; 3\,T_i\,\rho_i \times 100^2 \quad [\mathrm{g\,m^{-2}}]} \tag{8}$$

The 100^2 converts $\mathrm{cm^2}$ to $\mathrm{m^2}$ when T_i is in cm and ρ_i in $\mathrm{g\,cm^{-3}}$. (The "$\mathrm{g\,cm^{-2}}$" in the ATBD Table 2 header is a typographical error; calcite at 2.71 is plainly $\mathrm{g\,cm^{-3}}$.) The same factor 3 multiplies the quartz/feldspar mass in Eq. (9), so it cancels identically in the final mass ratio MR_i; only the absolute areal masses carry it. Appendix C gives the step-by-step derivation, the implicit assumptions, and an independent cross-check based on the Cauchy chord.

6 Link 5: The invisible matrix, quartz and feldspar

This is the EMIT-specific innovation. Quartz and feldspar dominate most soils by mass but produce no VSWIR features, so their optical path cannot be obtained from any band depth. The argument instead is that the *same* photons that encountered the detected minerals also random-walked through the quartz/feldspar matrix. If the average detected mineral underwent $\bar{m} = \frac{1}{j}\sum_k^j m_k$ scattering events, and n is the ratio of the number of mineral grains to the number of quartz/feldspar grains (a free parameter of the algorithm), then the quartz penetration depth follows the same diffusion law with the quartz diameter D_q scaled by the relative encounter frequency:

$$\boxed{T_q \;=\; \frac{\pi}{4}\,\frac{D_q}{n}\left(\frac{\sum_k^j m_k}{2j}\right)^{1/2}, \qquad M_q \;=\; 3\,T_q\,\rho_q \times 100^2} \tag{9}$$

Critically, D_q is *not* assumed: it is the median grain size predicted pixel by pixel by the random-forest model of Step 2 (Okin et al., 2025). This is why Step 2 exists, and why the ATBD's uncertainty budget is dominated by grain-size error: the quoted two-sigma envelope of $-49.5/+80.0\ \mu m$ is propagated through Eq. (9) as a systematic perturbation to produce the delivered 95% confidence intervals.

The final per-pixel product is the relative mass fraction

$$\boxed{MR_i \;=\; \frac{M_i}{\sum_k^j M_k + M_q}} \tag{10}$$

with the grain-size breakouts (goethite-nano/fine/med/large, etc.) summed into the ten delivered mineral classes, and the vegetation correction $MR_i^c = MR_i/f_s$ applied afterward using the Step 1 soil fraction f_s (pixels with $f_s < 0.65$ are masked).

6.1 Reading the subscripts: one soil, parallel pure columns

Everything carrying subscript i is tied to mineral i *and* to the wavelength of its diagnostic band (the "Reflectance Wavelength" column of Table 1): k_i is the continuum-removed absorption coefficient at that wavelength, m_i counts the photon's encounters with grains of mineral i, D_i is their diameter, and Eq. (7) then computes the RMS depth of a walk of m_i steps of size D_i, which is literally the penetration depth in a fictitious medium made of pure mineral i. The structure of the chain is therefore to treat each detected mineral as if it formed its own pure granular column, infer from its band depth how many scatterings occurred in that column, compute the column's depth and mass,

and ratio the columns at the end (Eq. (10)), much as one decomposes a mixed atmosphere into partial columns of its constituent gases. Physically there is one soil and one photon population: the photon walks through the real mixed medium (mostly quartz), and m_i is the subset of its encounters that involved mineral i, all extracted from the same measured spectrum. The pure column is bookkeeping, not a layer.

In Eq. (9), j is the number of minerals detected in the pixel and $k = 1, \ldots, j$ indexes them. The quantity $\bar{m} = \frac{1}{j}\sum_k m_k$ is the arithmetic mean encounter count over the detected species, what the ATBD calls the number of scattering events "the average detected mineral underwent". For a pixel with hematite ($m = 0.08$), kaolinite ($m = 0.31$), and calcite ($m = 0.02$): $j = 3$ and $\bar{m} = 0.137$. Since $\sum_k m_k/2j = \bar{m}/2$, Eq. (9) is Eq. (7) applied verbatim to the matrix with $m \to \bar{m}$ and $D \to D_q/n$. Each m_k is measurable because it derives from the *depth* of an observed band (the number of bands only sets j); m_q, the number of matrix encounters, is not measurable, because quartz and feldspar have no diagnostic band from which to start the chain of Eqs. (5)–(6). Physically m_q is the largest encounter count of all in a quartz-dominated soil; it is simply invisible. $\bar{m}$ is the observable proxy for it, and n is the exchange rate between the two.

6.2 The matrix is wavelength dependent

Because each mineral's quantities live at its own band wavelength, the composition of the "matrix" differs from band to band. At the 0.75 μm hematite band, kaolinite and calcite are essentially conservative scatterers and, together with quartz, they constitute the medium in which the hematite-absorbing photons random-walk; their Table 1 grain sizes (2 and 20 μm) differ enormously from the 100 μm quartz, so they materially change the step statistics and path-length distribution at that wavelength. Conversely, at 2.1 μm the hematite grains are (weakly absorbing) scatterers in the medium hosting the kaolinite band. In the Hapke picture this is automatic: the mixture albedo at any wavelength is the cross-section-weighted sum over *all* species, Eq. (3), and species that do not absorb at λ contribute with $w_k(\lambda) \approx 1$. The chain of Eqs. (5)–(9) contains no such term: T_{hem} depends on m_{hem} and D_{hem} only, with no knowledge of the grains that shaped the walk.

Two consequences follow. First, the other detected minerals must *not* be added to T_q even though they scatter at the hematite wavelength: they receive their own columns from their own bands, and M_q is reserved for species without a feature anywhere in the EMIT range; adding them would double count them in the denominator of Eq. (10). Second, averaging encounter counts obtained at different wavelengths into $\bar{m}$ is legitimate only under geometric optics (scattering, hence the walk geometry, is wavelength independent) and weak absorption (absorption does not truncate the walk). The second condition fails for hematite at 0.75 μm, so $\bar{m}$ mixes photon populations of different walk extent. The albedo-space inversion of Appendix D handles the wavelength dependence of the matrix by construction: every F_k is a fraction of the total cross-section at its own band, and $F_q = 1 - \sum_k F_k$ is the true remainder.

6.3 Reconstructing Eq. (9) from the encounter-share logic

The ATBD does not derive Eq. (9). Its stated logic, that the photons which produced the mineral bands also walked among matrix grains, nevertheless implies a specific derivation, and comparing it with the written equation locates three departures (Appendix F gives the details). Encounter counts are proportional to number density times geometric cross-section, so the matrix encounters relate to the total mineral encounters through the cross-section ratio,

$$m_q \;=\; \frac{1}{n_A}\sum_{k=1}^{j} m_k, \qquad n_A \;\equiv\; \frac{\sum_k N_k\,\sigma_k}{N_q\,\sigma_q} \;=\; n\,\frac{\bar{D}^2_{\min}}{D_q^2}, \tag{11}$$

where $n = N_{\min}/N_q$ is the grain *number* ratio of the ATBD's definition and $\bar{D}^2_{\min}$ is the number-weighted mean square mineral diameter. Applying Eq. (7) to the matrix encounters gives

$$T_q^{\text{derived}} = \frac{\pi}{4} D_q \sqrt{\frac{\sum_k m_k}{2\, n_A}} \qquad \text{versus} \qquad T_q^{\text{ATBD}} = \frac{\pi}{4} \frac{D_q}{n} \sqrt{\frac{\sum_k m_k}{2j}}. \tag{12}$$

The departures are: (i) *sum versus mean*, a factor $\sqrt{j}$ in T_q that varies from pixel to pixel and therefore does not cancel in MR_i; (ii) n outside the square root instead of inside, which for a global constant is an empty distinction, since $(D_q/n)\sqrt{\bar{m}/2} = D_q\sqrt{\bar{m}/2n^2}$, but becomes a real one as soon as n is given physical meaning or made pixel dependent; (iii) the cross-section factor $D_q^2/\bar{D}^2_{\min}$ is absent, silently absorbed into the calibrated n. Finally, D_q is the median grain diameter of the pixel's bulk texture from Step 2, the interpolated 50th percentile of the predicted clay/silt/sand distribution, and the only pixel-varying grain size in the chain; it is a proxy for the matrix grain size that is clean in sandy soils and pulled toward the clay-sized detected minerals elsewhere.

7 Mineral group metadata

Table 1 reproduces the fixed per-group parameters used in Eqs. (5)–(8) (ATBD Table 2). These are held constant globally within each group, a stated assumption and limitation of the approach.

Mineral group	k (`delta_k`) (cm^{-1})	ρ (g cm^{-3})	Band center (μm)	D (μm)
calcite	10.51	2.71	2.18	20
chlorite	76.96	2.80	2.14	11
dolomite	10.51	2.84	2.17	20
goethite-nano	1000	3.80	0.76	0.05
goethite-fine	5120	3.80	0.76	1
goethite-med	5120	3.80	0.76	2
goethite-large	5120	3.80	0.76	5
gypsum-fine	50	2.36	2.09	5
gypsum-coarse	50	2.36	2.09	20
hematite-nano	1304	5.30	0.75	0.05
hematite-fine	6400	5.30	0.75	1
hematite-med	6400	5.30	0.75	2
hematite-large	6400	5.30	0.75	5
illite+muscovite	91.2	2.78	2.14	2
kaolinite	181	2.60	2.11	2
montmorillonite	100	2.35	2.12	2
vermiculite	100	2.50	2.15	10

Table 1: Mineral groups and metadata, transcribed from ATBD Table 2.

8 A worked numerical example

Take a pixel where Tetracorder reports hematite-fine with $BD = 0.10$, and Step 2 gives $D_q = 100\ \mu m$. From Table 1: $k = 6400\ \text{cm}^{-1}$, $D = 1\ \mu m$, $\rho = 5.3\ \text{g cm}^{-3}$.

$$\bar{d} = \frac{\ln(0.9)}{-6400\ \mathrm{cm}^{-1}} = 1.65 \times 10^{-5}\ \mathrm{cm} = 0.165\ \mu m, \tag{13}$$

$$m = \frac{0.165}{2 \times 1} = 0.082, \tag{14}$$

$$T = \frac{\pi}{4} \times 1\ \mu m \times \left(\frac{0.082}{2}\right)^{1/2} = 0.16\ \mu m, \tag{15}$$

$$M = 3 \times 1.6 \times 10^{-5}\ \mathrm{cm} \times 5.3 \times 10^{4} \approx 2.5\ \mathrm{g\,m}^{-2}. \tag{16}$$

For the quartz/feldspar matrix with $n = 1$ and $j = 1$:

$$T_q = \frac{\pi}{4} \times 100\ \mu m \times (0.041)^{1/2} \approx 16\ \mu m, \tag{17}$$

$$M_q = 3 \times 1.6 \times 10^{-3}\ \mathrm{cm} \times 2.65 \times 10^{4} \approx 127\ \mathrm{g\,m}^{-2}, \tag{18}$$

hence

$$MR_{\mathrm{hem}} = \frac{2.5}{2.5 + 127} \approx 1.9\%. \tag{19}$$

A 10% band depth maps to roughly a 2% mass fraction. The example makes the leverage of each parameter visible: the nonlinearity of $\ln(1 - BD)$ matters little at small depths, but $M_q \propto D_q/n$ rescales every retrieved fraction directly, which is exactly why grain size carries the error budget and why n is an honest free parameter rather than a derived quantity.

9 Assumptions and limitations

The per-group k, ρ, and D are held constant globally (the ATBD flags this explicitly), so within-class mineralogical variability (e.g. Fe substitution in clays) aliases into mass error. The Beer–Lambert linearization degrades for very strong absorbers, which is partly why the iron oxides receive separate nano/fine/med/large classes with different effective k and D, following the grain-size-dependent band saturation behavior documented by Clark and Roush (1984). The quartz/feldspar inference assumes the detected photons and the matrix photons share one diffusion geometry, and MR_i is relative only to the retrieved minerals plus quartz/feldspar, not to total soil mass including amorphous or undetected phases. Finally, the band depths inherit Tetracorder's library-matching assumptions: each depth is measured after a least-squares fit of continuum-removed library spectra (the "fit" and the 1 μm and 2 μm groups of Clark et al. 2003), so library mismatch propagates silently, a point the ATBD lists among its unquantified uncertainties.

10 A critical assessment: ranking the assumptions

Having worked through every link of the chain, we can rank the algorithm's assumptions from weakest and most vulnerable to most defensible. For each we state the leverage (how errors propagate to the delivered product) and the improvement path. Item 0 concerns the framework itself; items 1 through 5 concern its ingredients.

0. The closure of the transport problem (the hybrid framework; weakest of all). The framework is a hybrid: its authority derives from Hapke's radiative transfer theory (Eq. (2)), but the mass chain of Eqs. (5)–(10) is built from geometric-optics photon bookkeeping and a diffusion argument. Is the hybrid physically consistent at the fundamental level? In one respect, yes, and more deeply than it appears: the radiative transfer equation and the photon random walk are two descriptions of the same stochastic transport process. The formal solution of the RTE is an average over photon paths,

$$R = \int_0^\infty P(\ell)\, e^{-k\ell}\, d\ell, \tag{20}$$

with $P(\ell)$ the path-length distribution of the conservative scattering problem; two-stream and Eddington closures are the P_1 (diffusion) approximation of the RTE; and diffusion is the long-time limit of a random walk. Hapke's own microphysics is geometric optics (the slab model of Appendix A). The two halves share both microphysics and the underlying process.

The genuine seam is the *closure*. Hapke performs a proper ensemble average over the path distribution, which matters because $P(\ell)$ for a semi-infinite diffusive medium is heavy-tailed ($P(\ell) \sim \ell^{-3/2}$; the mean return path formally diverges). The Laplace transform of that tail produces the signature nonlinearity of reflectance spectroscopy: the albedo deficit scales as $\gamma = \sqrt{1-w}$, so for a dilute absorber of cross-section fraction F in a bright matrix the band depth grows as $BD \propto \sqrt{F\,kD}$, not linearly. The chain of Eqs. (5)–(7) instead follows a single mean photon: Beer–Lambert on an effective path, $BD \approx k\bar{d}$, a first-moment truncation of Eq. (20). The $\bar{d}$ it retrieves is not the ensemble mean path (which diverges); it is the path of a fictitious equivalent single-transit absorber.

The consequence is a clash of scaling laws, summarized in Table 2. The EMIT chain gives $M_i \propto \sqrt{D_i \bar{d}_i} \propto \sqrt{BD}$. Single-scattering physics (dark matrices, saturated bands) gives $F \propto BD$. Diffusive multiple scattering in a bright matrix, the typical regime for a few-percent absorber in quartz sand, gives $F \propto BD^2$. Across a band-depth range of 0.05 to 0.5, the three laws spread retrieved masses by factors of roughly 3, 10, and 100 respectively. The framework's implied exponent of one half matches neither limiting regime and is the most compressive of the three. Unlike the order-unity constants of Appendices A–C, this does not cancel in MR_i: it redistributes mass among minerals of different band depths and biases the contrast between abundant and trace species.

Framework	Closure	Mass–band-depth scaling
EMIT L3 chain	mean-path Beer–Lambert	$M \propto BD^{1/2}$
Single scattering	exact low-order limit	$F \propto BD$
Hapke / diffusive	ensemble over heavy-tailed $P(\ell)$	$F \propto BD^2$

Table 2: Implied scaling of retrieved abundance with band depth under three closures of the same transport problem.

Three effects have kept this from being catastrophic in practice. The Tetracorder band depths are measured relative to pure library samples and the group metadata were tuned against laboratory mixtures, anchoring the absolute level at calibration conditions; the wrong exponent bites mainly in extrapolation across the dynamic range. Continuum absorption of the matrix truncates the heavy tail of $P(\ell)$, pulling the true exponent back from 2 toward 1. And the free parameter n silently absorbs part of the residual. None of this is a derivation; it is calibration covering for a closure inconsistency. A self-consistent alternative exists within the Hapke framework itself and was prescribed by Clark and Roush (1984): invert reflectance to single-scattering albedo first, then

unmix linearly in albedo space. Appendix D gives the algorithm; Appendix E outlines a benchmark study to quantify the closure error directly.

At the deepest level one may also question the incoherent transport assumption itself: both frameworks use far-field single-particle properties at a filling factor near 0.5, formally outside the RTE's domain of validity. This is the dense-media radiative transfer problem (e.g., the quasicrystalline approximation of Tsang et al., 2000). For VSWIR soil grains the size parameter is in the hundreds, so the coherent corrections that dominate at microwave wavelengths are second-order here; the leading dense-packing effects are geometric (porosity and shadowing, treated by Hapke's later porosity factor K). The closure inconsistency above is the first-order, actionable concern.

1. The free parameter $n \equiv 1$. The ATBD defines n as the ratio of mineral grain number to quartz/feldspar grain number, and in practice it is a single global constant. Taken literally it is off by orders of magnitude: a soil with 10% clay by mass at $D_c = 2\ \mu m$ against 100 μm quartz has a *number* ratio near 10^4, not 1, because number scales as $M/(\rho D^3)$. Unlike the geometric conventions, n does not cancel: it enters $M_q \propto D_q/n$ linearly and alone, so it directly sets the mineral-versus-matrix partition, i.e. the absolute level of every delivered mass fraction. It is not varied in the uncertainty budget. In truth n is not a physical quantity at all; it is the calibration dial that absorbs every discarded order-unity factor in the chain (Appendices A–C). Improvement path: derive an effective n per pixel from the Step 2 texture prediction itself (clay fraction plus grain sizes yield the number ratio), or eliminate it by reformulating Link 5 as a Hapke intimate-mixture inversion of the single-scattering albedo, where the quartz fraction enters through Eq. (3) with no free grain-count parameter.

2. Intimate mixture of equant, free grains; no coatings. The mass conversion assumes every mineral exists as independent spherical grains of diameter D_i. In real desert soils, hematite and goethite exist substantially as nanophase *coatings and rinds on quartz grains*, which is optically a thin-film problem, not a volume-scattering problem: the mass-to-band-depth relation is completely different. The algorithm knows this at the detection level (the Tetracorder library includes "thin film" and "coated" entries) but pushes them through the same sphere-chord mass conversion with an effective D of 0.05 to 1 μm. Worse, at $D = 0.05\ \mu m$ the size parameter is below 1 and the entire geometric-optics chain (chords, per-transit Beer–Lambert, Fresnel branching) is formally invalid; the nano-class k and D are empirical patches. Since the iron oxides drive the mission's headline radiative-forcing objective, this vulnerability sits exactly where the science payoff is. Improvement path: a separate forward-model pathway for coated morphologies, validated against laboratory analogs of coated quartz.

3. Globally constant per-group metadata (k, ρ, D). One absorption coefficient and one diameter per class, worldwide (Table 1). Within-class variability is large: Al-substitution and crystallinity change iron oxide k by factors of several; clay band strengths vary with composition and hydration. The saving grace is the square root: $M_i \propto \rho_i \sqrt{D_i/k_i}$, so a factor-of-2 error in k gives only 40% in mass, and part of it cancels in the ratio. But the discrete quantization of iron oxide grain size into four bins, with k jumping from 1000 to 5120 between nano and fine, is crude, and EMIT's own data contain the remedy: band position, width, and asymmetry carry continuous grain-size and crystallinity information that the pipeline discards after Tetracorder matching. Continuous spectral-shape inversion of the Fe-oxide bands is an obvious upgrade.

4. The heuristic transport bookkeeping. Beer–Lambert linearization, the $2D$ chord, the $\pi D/4$ step, the factor 3. The uniform constants cancel in MR_i (Appendices A–C), so this ranks lower-risk than the items above. What does not cancel is structural: the $\sqrt{m}$ weighting, versus the linear-m weighting of the exact straight-ray bookkeeping (Eq. (41)), redistributes mass among coexisting minerals, compressing strong absorbers relative to weak ones; and the weak-absorption limit fails precisely at the saturated Fe-oxide band centers. Mitigations exist (calibration lineage, ratio structure), but the model has not, to our knowledge, been benchmarked head-to-head against full nonlinear Hapke unmixing or Monte Carlo radiative transfer on laboratory intimate mixtures of known mass fractions in the multi-mineral regime. That benchmark is a well-posed, self-contained study.

5. All undetected mass is quartz/feldspar with the Step 2 grain size (most justified). The assumption is physically sound for dust-source soils, the random-forest grain size is validated to roughly 8–12%, and its uncertainty is the one error source honestly propagated into the delivered confidence intervals. Residual risks are second-order: amorphous phases and volcanic glass are also spectrally bland but are not quartz, and the machine-learning model extrapolates poorly outside its training domain. This is what a justified assumption looks like: explicit, validated, propagated.

Beyond the algorithm: the surface-versus-reservoir question. One assumption sits above the L3 algorithm but may matter most for the mission science: the measurement weighs the top tens of microns of soil (the worked example gives $T_q \approx 16\ \mu m$), while saltation mobilizes millimeters to centimeters. Crusts, desert varnish, and coatings make the optical surface systematically different from the erodible reservoir. Combining items 1 and 2 above (a coating-aware, free-parameter-free intimate mixture inversion) with this surface-versus-reservoir question defines a coherent space for further investigation.

11 Logic gaps and internal inconsistencies

Section 10 ranked the assumptions by vulnerability. This section collects, in one place, the points at which the ATBD's chain is internally inconsistent or where a stated premise is contradicted by the algorithm's own equations or parameter values. They are numbered for reference, not by severity.

1. **Half Hapke, half ray tracing, and neither carried through.** The ATBD names Hapke (1981) and Clark and Roush (1984) as its theoretical basis, but no equation from either is used: the reflectance-to-albedo relation, the H-functions, and the intimate mixing rule (Eqs. (2)–(3)) never appear in the chain. The chain is instead single-photon ray-tracing bookkeeping (Beer–Lambert along a path, discrete chords, a random walk), yet no ray tracing is performed either: every step substitutes a mean value (mean chord, mean step, RMS depth) for the ensemble average that the physics requires. Radiative transfer is invoked but not applied; ray tracing is implied but not executed. Item 0 of Section 10 quantifies the resulting closure error.

2. **Band depth as transmittance.** The ATBD writes $BD = 1 - t$ with $t = e^{-k\bar{d}}$, identifying the reflectance ratio R_b/R_c with a single-pass transmittance. Clark and Roush (1984), cited for the definition of band depth, show that in the multiple-scattering regime the reflectance deficit scales as $\sqrt{1-w}$, not linearly in absorption. The identification is the first-moment truncation of Eq. (20) and is the root of the exponent clash in Table 2.

3. **A random walk of less than one step.** m_i is called "the number of scattering events" and is then inserted into the diffusion formula $T \propto \sqrt{m}$, whose derivation (Appendix B) requires $m \gg 1$. With the Table 1 parameters, typical values are $m_i \sim 0.02$ to 0.5 (the worked example gives 0.08). The chain thus applies the asymptotic law of a many-step walk to fractional step counts. As an expectation value the construction is definable, but the $\sqrt{m}$ scaling has no derivation in this regime.

4. **Weak absorption assumed, strong absorption tabulated.** The Beer–Lambert linearization (Link 1), the geometric series behind the $2D$ chord (Appendix A), and the wavelength-independence of the walk (Section 6.2) all require $kD \ll 1$. Table 1 lists $k = 6400$ cm^{-1} with D up to 5 μm for hematite, i.e. kD up to 3, and single-transit transmittance e^{-2kD} below 10^{-2}. The iron oxides, the minerals that matter most for radiative forcing, sit outside the regime in which the chain's own approximations hold.

5. **Parallel pure columns in an intimate mixture.** The ATBD introduces the framework by invoking intimate mixtures, in which multiple scattering couples all components (Section 6.1). The chain then computes each mineral's depth and mass in isolation, as if it formed a pure column unaffected by the grains around it, and sums the columns. The coupling that motivated the framework is absent from its equations.

6. $\bar{m}$ **averages incommensurable quantities.** The encounter counts entering $\bar{m}$ are obtained at different wavelengths (0.75 versus 2.1 μm), from photon populations whose walk extents differ because absorption differs (Section 6.2), and they scale with each mineral's abundance rather than estimating a common walk property. The arithmetic mean of such numbers has no clear physical meaning; the ATBD's own encounter-share logic implies the sum (Section 6.3), and the two differ by a pixel-dependent factor $\sqrt{j}$.

7. n**: defined as one thing, used as another.** n is defined as a grain *number* ratio, which for real soils ranges over six orders of magnitude, but it is held constant and enters as D_q/n outside the square root. The written form is consistent only with a cross-section (area) ratio, and only after absorbing the factor $D_q^2/\bar{D}_{\min}^2$ (Section 6.3, Appendix F). The parameter's name, its stated definition, and its algebraic role are three different things.

8. D_q**: a bulk median labeled as a quartz diameter.** D_q is the median of the whole predicted texture distribution, including the clay-sized fraction that is largely the detected minerals themselves. In clay-rich pixels the "quartz/feldspar grain diameter" is set mostly by non-quartz grains. Since $M_q \propto D_q$ is the most leveraged term in the product, the label matters.

9. **Geometric optics applied at size parameter below unity.** The nano classes ($D = 0.05$ μm at $\lambda = 0.75$ μm, $x \approx 0.4$) are pushed through chords, Snell refraction, Fresnel branching, and per-transit Beer–Lambert, none of which exist in the Rayleigh-to-resonance regime. Their k values (1000 to 1304 versus 5120 to 6400 for the same minerals at $D = 1$ μm) are effective fitting constants, not optical properties, contradicting the presentation of k as an absorption coefficient.

10. **Solid columns in a porous medium.** Eq. (8) weighs a solid column of mineral at bulk density with no filling factor, while the intimate-mixture theory it invokes is explicitly a theory of porous particulate media (and later Hapke versions carry a porosity factor K). The factor cancels in MR_i only because it is omitted uniformly; the absolute areal masses in g m^{-2} are

not physical. The "g cm^{-2}" density unit in Table 1 is a related sign of loose dimensional bookkeeping.

11. **The vegetation correction acts on a ratio.** Section 4 of the ATBD applies $MR_i^c = MR_i / f_s$. Division by the soil fraction is the correct areal-mixing correction for an *extensive* quantity such as a band depth or an areal mass. MR_i is an intensive ratio in which the vegetation dilution of the numerator (through BD_i) and of the denominator (through $\bar{m}$ in M_q) largely cancels; dividing the ratio by f_s can double correct and can produce $MR_i^c > 1$. Whether the delivered code corrects band depths or mass fractions should be verified.

12. **An uncertainty budget that omits the structural terms.** The delivered confidence interval perturbs D_q only. Not propagated: n, the per-class k_i and D_i, the sum-versus-mean choice, the closure exponent, or the coating morphology. The ATBD argues that random measurement errors average down over 0.5° cells; the errors listed here are systematic and spatially coherent, and do not.

None of these items is fatal to the product as a calibrated, relative map of mineral abundance, which is what the delivered 0.5° fields are used as. Together they do mean that the chain should be read as an empirical transfer function anchored at its calibration conditions rather than as the derivation from first principles that its presentation suggests, and that the albedo-space formulation of Appendix D would remove items 1, 2, 5, 6, 7, and 10 at once.

A Appendix: The effective chord and the factor $2D_i$ in Eq. (6)

Equation (6) converts the total in-grain path $\bar{d}_i$ into a number of grain encounters by asserting that each encounter contributes an *effective chord* of $2D_i$ of absorbing path. This appendix derives that factor from Hapke's slab model of particle transmission and shows why the exact value does not affect the delivered mass fractions.

A.1 The geometric baseline: Cauchy's mean chord

For a convex body under uniform, isotropic illumination, the mean straight-line chord is the Cauchy result

$$\langle \ell \rangle \;=\; \frac{4V}{S}, \tag{21}$$

which for a sphere of diameter D gives

$$\langle \ell \rangle \;=\; \frac{4\,(\pi D^3/6)}{\pi D^2} \;=\; \frac{2}{3}D \;\approx\; 0.67\,D. \tag{22}$$

If photons passed straight through grains, the denominator of Eq. (6) would be $2D_i/3$. The factor-of-three difference between this and the adopted $2D_i$ arises from two refractive effects that lengthen the optical path relative to the geometric chord.

A.2 Effect 1: Refraction weights the long chords

Soil minerals have real refractive index $n \approx 1.5$. Snell's law bends entering rays toward the surface normal, concentrating the internal ray bundle toward central, near-diameter chords and de-weighting the grazing short ones. Hapke's slab-model estimate of the mean single-transit internal path is

$$\langle D \rangle \;=\; \frac{2}{3}\left[n^2 - \frac{1}{n}\left(n^2-1\right)^{3/2} \right] D \;\approx\; 0.9\,D \qquad (n = 1.5), \tag{23}$$

already 35% longer than the Cauchy chord.

Derivation of Eq. (23)

Let the sphere have radius $a = D/2$ and be illuminated by a uniform parallel beam. A ray with impact parameter b (perpendicular distance from the central axis) strikes the surface at local incidence angle θ, with $b = a\sin\theta$. The fraction of the beam intercepted between b and $b + db$ is the annular area over the total cross-section,

$$dP \;=\; \frac{2\pi b\, db}{\pi a^2} \;=\; 2\sin\theta\cos\theta\, d\theta, \qquad \theta \in [0, \pi/2]. \tag{24}$$

This weighting already favors oblique incidence (it peaks at $\theta = 45°$), which is why the unrefracted average falls below the diameter.

At entry, Snell's law gives the internal angle θ' via $\sin\theta = n\sin\theta'$. Inside the sphere the ray travels along a chord; since the local normal is radial, a chord making angle θ' with the normal has length

$$\ell(\theta') \;=\; 2a\cos\theta' \;=\; D\cos\theta', \tag{25}$$

as seen by dropping a perpendicular from the center onto the chord (the half-chord is $a\cos\theta'$). Without refraction ($\theta' = \theta$), averaging $D\cos\theta$ over Eq. (24) gives $\int_0^{\pi/2} D\cos\theta \cdot 2\sin\theta\cos\theta\, d\theta = \frac{2}{3}D$, recovering the Cauchy chord of Eq. (22). Refraction bends each ray toward the normal, $\theta' < \theta$, so $\cos\theta' > \cos\theta$: every internal chord is longer than its unrefracted counterpart.

Averaging Eq. (25) over the beam, with $\cos\theta' = \sqrt{1 - \sin^2\theta/n^2}$,

$$\langle D\rangle \;=\; \int_0^{\pi/2} D\sqrt{1 - \frac{\sin^2\theta}{n^2}}\; 2\sin\theta\cos\theta\, d\theta \;=\; 2D\int_0^1 u\sqrt{1 - \frac{u^2}{n^2}}\, du, \tag{26}$$

using $u = \sin\theta$. Substituting $v = 1 - u^2/n^2$, so that $u\, du = -\frac{n^2}{2}\, dv$ with limits $v : 1 \to 1 - 1/n^2$,

$$\langle D\rangle \;=\; 2D\cdot\frac{n^2}{2}\int_{1-1/n^2}^{1} v^{1/2}\, dv \;=\; \frac{2}{3}D\, n^2\left[1 - \left(1 - \frac{1}{n^2}\right)^{3/2}\right]. \tag{27}$$

Finally, $n^2\left(1 - 1/n^2\right)^{3/2} = (n^2 - 1)^{3/2}/n$, which yields Eq. (23) exactly.

Two limits check the result. As $n \to 1$, $\langle D\rangle \to \frac{2}{3}D$, the Cauchy chord, as it must. As $n \to \infty$, expanding $(1 - 1/n^2)^{3/2} \approx 1 - \frac{3}{2n^2}$ gives $\langle D\rangle \to D$: strong refraction bends every entering ray toward the radial direction, so all rays traverse near-diametral chords and the mean path saturates at the diameter. Physically, refraction does two things at once: it lengthens each chord and it redistributes the beam flux into the longer, central chords.

The derivation assumes geometric optics ($D \gg \lambda$), a homogeneous sphere, and a single transit (internal reflections are handled separately below). It also weights all entering rays equally, ignoring the angular dependence of the Fresnel transmission coefficient; in reality transmission drops sharply at grazing incidence, further de-weighting the short chords and pushing $\langle D\rangle$ slightly closer to D. A formal subtlety: what matters physically in Hapke's internal transmission factor Θ is the mean transmittance $\langle e^{-\alpha\ell}\rangle$, whereas the model uses $e^{-\alpha\langle\ell\rangle}$. By Jensen's inequality $\langle e^{-\alpha\ell}\rangle \geq e^{-\alpha\langle\ell\rangle}$, so the equivalent-slab substitution slightly overestimates absorption; the two agree to first order in $\alpha\ell$, which is again the weak-absorption limit the EMIT chain relies on.

A.3 Effect 2: Internal reflections multiply the transits

A ray arriving at the exit interface from inside the grain sees the critical angle, $\theta_c = \arcsin(1/n) \approx 42°$ for $n = 1.5$. All internal rays beyond θ_c are totally internally reflected, and even within the critical cone the Fresnel reflectivity is nonzero. Averaged over the internal hemisphere, Hapke's approximation for the internal surface reflectivity is

$$S_i \;\approx\; 1.014 - \frac{4}{n(n+1)^2} \;\approx\; 0.59 \qquad (n = 1.5). \tag{28}$$

After each internal segment the photon therefore has a probability $S_i \approx 0.6$ of reflecting back into the grain and traversing another segment. In the weak-absorption limit the segment count forms a geometric series with expectation $1/(1 - S_i)$, and the expected total in-grain path per encounter is

$$\ell_{\text{eff}} \;=\; \frac{\langle D \rangle}{1 - S_i} \;\approx\; \frac{0.9\,D}{0.41} \;\approx\; 2.2\,D. \tag{29}$$

This is the origin of the $2D_i$ in Eq. (6): one refraction-weighted transit of $\approx 0.9D$, multiplied by an expected 2 to 2.5 internal segments before escape. The adopted value of exactly 2 is a rounded convention (Clark et al. 2024, Section 11.4), but it is quantitatively consistent with Hapke's internal-transmission model for a weakly absorbing grain of ordinary silicate index. The same structure appears in Hapke's internal transmission factor for the particle single-scattering albedo, where the recursion $\Theta = e^{-\alpha\langle D\rangle}/(1 - S_i\,\Theta)$-type terms encode precisely this multiple-internal-bounce lengthening.

Note also the meaning of m_i that follows: it is an expectation value, not an integer count. In practice $m_i < 1$ is common, to be read as the probability that the average detected photon transited a grain of mineral i during its random walk through the soil.

A.4 Caveats

Strong absorbers. The geometric series in Eq. (29) assumes absorption is weak enough that later internal segments are not attenuated. At the band center of a strong absorber (hematite at 0.75 μm, where $k\,D \sim 1$), the second and third segments are partially absorbed away, so the true effective chord drops below $2D$ and Eq. (6) overestimates the path per encounter. This is the same weak-absorption linearization that underlies Link 1 (Eq. (5)), so the approximation is at least internally consistent.

Insensitivity of the final product. The delivered mass fractions are insensitive to the exact chord constant. Replace $2D_i$ by cD_i for any constant c applied uniformly to all minerals: every m_i scales by $2/c$, every penetration depth $T \propto \sqrt{m}$ scales by $\sqrt{2/c}$, and hence every areal mass M_i and M_q carries the same factor $\sqrt{2/c}$, which cancels identically in

$$MR_i \;=\; \frac{M_i}{\sum_k M_k + M_q}. \tag{30}$$

The choice of 2 versus 0.67 or 0.9 affects the absolute areal masses in $\mathrm{g\,m^{-2}}$ but not the relative mass fractions that constitute the L3 product. What would matter is applying *different* chord conventions to different minerals, which the algorithm does not do.

B Appendix: The random walk and the penetration depth of Eq. (7)

Equation (7),

$$T_i \;=\; \frac{\pi}{4}\, D_i \left(\frac{m_i}{2}\right)^{1/2}, \tag{31}$$

converts a *number of scattering events* m_i into a *depth below the surface* T_i. The physics is that a multiply scattered photon does not propagate ballistically into the soil; it random-walks, so the depth it probes grows as the square root of the number of steps. The equation is exactly the RMS vertical displacement of a random walk with a particular choice of step length. This appendix derives it, explains the physical meaning of the step length s, and discusses how the branching between external reflection and refraction at each grain encounter is handled.

B.1 Derivation

Step 1: the step length $s = \pi D/4$. Between successive scattering events in a closely packed granular medium, a photon is displaced by a distance of order one grain diameter. The specific value adopted is the mean chord of a circular cross-section of diameter D. For a circle of radius $a = D/2$, a straight path at perpendicular offset x from the center has chord length $2\sqrt{a^2 - x^2}$; averaging uniformly over offsets,

$$s \;=\; \frac{1}{2a}\int_{-a}^{a} 2\sqrt{a^2 - x^2}\, dx \;=\; \frac{\pi a^2}{2a} \;=\; \frac{\pi}{4} D \;\approx\; 0.785\, D. \tag{32}$$

So $\pi D/4$ is the average distance across a grain's circular silhouette for a randomly positioned traversal, taken as the characteristic event-to-event hop of the walk.

Step 2: variance addition. Let the photon take m steps of length s, with direction randomized at every scattering event. The vertical (depth) component of step k is $z_k = s\cos\phi_k$. Because successive directions are uncorrelated, mean displacements cancel, $\langle z\rangle = 0$, but variances add:

$$\langle z^2\rangle \;=\; \sum_{k=1}^{m} \langle z_k^2\rangle \;=\; m\, s^2\, \langle \cos^2\phi\rangle. \tag{33}$$

This is the core of the result and the origin of the $\sqrt{m}$: direction randomization at each grain encounter destroys ballistic penetration and replaces it with diffusion.

Step 3: the projection factor. Equation (7) treats the walk as isotropic in a vertical plane: ϕ uniform on $[0, 2\pi)$, so $\langle \cos^2\phi\rangle = 1/2$ and

$$T \;=\; \sqrt{\langle z^2\rangle} \;=\; s\sqrt{\frac{m}{2}} \;=\; \frac{\pi}{4} D \left(\frac{m}{2}\right)^{1/2}, \tag{34}$$

which is Eq. (7) exactly. A fully three-dimensional isotropic walk would give $\langle \cos^2\theta\rangle = 1/3$ and $T = 0.577\, D\sqrt{m}$ versus the adopted $0.555\, D\sqrt{m}$; the 4% difference is far below the model's other uncertainties.

B.2 Why a two-dimensional silhouette, when reality is 3D?

The step length of Eq. (32) is the mean chord of the grain's *two-dimensional silhouette*, not of the three-dimensional grain. This deserves comment, both because the two averages differ and because the choice turns out to have a clean 3D interpretation.

2D and 3D Cauchy chords. For a convex body traversed by uniform random straight lines (μ-randomness), the Cauchy mean-chord formulas are

$$\langle \ell \rangle_{2\mathrm{D}} = \frac{\pi A}{P}, \qquad \langle \ell \rangle_{3\mathrm{D}} = \frac{4V}{S}, \tag{35}$$

with A, P the area and perimeter of a plane figure and V, S the volume and surface area of a solid. For a disk of diameter D this gives $\pi D/4 \approx 0.785\, D$ (Eq. (32)); for a sphere it gives $2D/3 \approx 0.67\, D$ (Eq. (22)). If one asked for the expected chord inside a *sphere* under isotropic illumination, the answer would be $2D/3$, not $\pi D/4$. Note also that "mean chord" is a Bertrand-paradox-type quantity: it depends on the measure used to generate the random lines. The values above hold for uniform random lines; other randomizations differ (e.g., chords of a circle drawn between two uniform points on the circumference average $2D/\pi \approx 0.64\, D$). The derivation in Eq. (32) uses uniform parallel offsets, which is the μ-randomness measure.

The projection interpretation of $\pi/4$. The observable computed by Eq. (7) is a depth, a one-dimensional quantity, and there are two equally legitimate routes from a 3D trajectory to a 1D variance. The convention adopted here takes the route through a plane. Consider the physically honest picture: a 3D random walk with steps of length $\approx D$ (one center-to-center hop between touching grains), projected onto any vertical plane. No depth information is lost, since the vertical axis lies in the plane. A 3D step of isotropic direction projects onto the plane with length $D \sin\theta_n$, where θ_n is the angle from the plane's normal, and the mean projection factor is

$$\langle \sin\theta_n \rangle \;=\; \int_0^{\pi} \sin\theta_n \, \frac{\sin\theta_n}{2}\, d\theta_n \;=\; \frac{\pi}{4}. \tag{36}$$

So $\pi D/4$ is *exactly* the mean projected step of a 3D isotropic walk of step D onto a plane; Eq. (7) is the 2D shadow of a 3D walk, and the shadow contains everything the depth calculation needs. That the same number is also the mean chord of the circular silhouette is no coincidence: both are averages of $\sqrt{1-x^2}$-type integrands over the same measure.

How good is the shadow route? Computed exactly, the vertical variance per step via the projection route is

$$D^2 \,\langle \sin^2\theta_n \rangle \,\langle \cos^2\phi \rangle \;=\; D^2 \cdot \frac{2}{3} \cdot \frac{1}{2} \;=\; \frac{D^2}{3}, \tag{37}$$

identical to the direct 3D result $D^2\langle \cos^2\theta \rangle = D^2/3$, as it must be. Equation (7) instead uses $\langle \sin\theta_n \rangle^2 = (\pi/4)^2$ in place of $\langle \sin^2\theta_n \rangle = 2/3$, a square-of-the-mean versus mean-of-the-square substitution. That is the entire content of the "2D versus 3D" discrepancy: $(\pi/4)^2/2 = 0.308$ versus $1/3 = 0.333$, an 8% difference in variance and 4% in depth.

The choice is therefore best understood in three layers. Physically, depth is a projection, so working in the plane containing the vertical loses nothing in principle. Practically, the planar bookkeeping with $\langle \ell_p \rangle = \pi D/4$ commits only the Jensen-type substitution above. Structurally, the model is deliberately built so the choice cannot matter: the 4%, like every other uniform

order-unity factor in the chain, cancels identically in the delivered mass fractions and is degenerate with the calibrated parameter n. The rigor is spent where the product is sensitive (grain diameters, absorption coefficients) and the cheapest defensible convention is used where it is not.

B.3 Physical meaning of the step length s

The step s must not be confused with either the mean free path of dilute-medium transport theory or the absorbing chord of Eq. (6). Figure 1 illustrates the distinction.

In a dilute medium (aerosols, cloud droplets), the mean free path is the free-flight distance through the host medium between two particle encounters, $l = 1/(N\sigma_{\text{ext}})$, and the particles are effectively points along the trajectory. In a soil with filling factor $\phi \approx 0.5$ this picture degrades. Applying the dilute formula with geometric cross-sections gives $l = 2D/(3\phi) \approx 1.3\,D$, already of order one grain diameter. More directly, the mean chord of the *pore phase* in a two-phase medium is $\frac{2D}{3} \cdot \frac{1-\phi}{\phi} \approx 0.7\,D$ at $\phi = 0.5$: the photon is almost never in open space for more than a fraction of a grain diameter. Moreover, a large fraction of scattering events are forward transmissions *through* grains, whose in-grain path is what does the absorbing (Eq. (6)).

A typical cycle of the walk is therefore: cross a pore gap of $\sim 0.7D$, then either reflect off a grain surface or refract through the grain interior (internal path $\sim 0.9D$ per transit, Eq. (23)), then emerge in a new direction and repeat. The step s of Eq. (7) is the *net displacement between successive scattering events*, lumping the pore gap and the grain transit into a single hop of order one grain diameter (panel A of Figure 1). The model thus tracks two distinct lengths from the same trajectory: the effective chord $2D_i$ of Eq. (6), the absorbing path accumulated *inside* grains, which feeds the Beer–Lambert bookkeeping; and the geometric step $\pi D/4$ of Eq. (7), the spatial displacement per event, which feeds the diffusion depth. Same grain, two different projections of the same trajectory.

B.4 External reflection versus refraction at each encounter

At every grain encounter two branches exist: external (surface) reflection, in which the photon never enters the grain, and refraction into the grain followed by zero or more internal reflections and eventual transmission out. The derivation above does not follow either branch individually; the rigorous theory averages over both with Fresnel probabilities, and the constants of Eqs. (6) and (7) inherit that average.

Branching probabilities. At the entry interface the split is governed by the Fresnel reflectivity at the local incidence angle. For $n = 1.5$ the normal-incidence reflectivity is $R(0) = \left(\frac{n-1}{n+1}\right)^2 = 0.04$, and averaged over the impact-parameter distribution of a sphere (Hapke's S_e), the external reflection probability is approximately 0.09: roughly one encounter in ten is an external bounce, and nine in ten the photon enters the grain. Once inside, the internal-wall reflectivity is much higher, $S_i \approx 0.59$ (Eq. (28)), because rays beyond the 42° critical angle are totally internally reflected; from outside there is no critical angle, which is why $S_e \ll S_i$. This asymmetry is the reason the effective chord of Eq. (6) is $\sim 2D$ rather than a single transit.

The expectation over the branch tree. Hapke's equivalent-slab expression for the particle single-scattering albedo is literally the probability ledger over all branches:

$$w \;=\; \underbrace{S_e}_{\text{external bounce}} \;+\; (1 - S_e)\,\frac{(1 - S_i)\,\Theta}{1 - S_i\,\Theta}, \qquad \Theta = e^{-\alpha\langle D\rangle}. \tag{38}$$

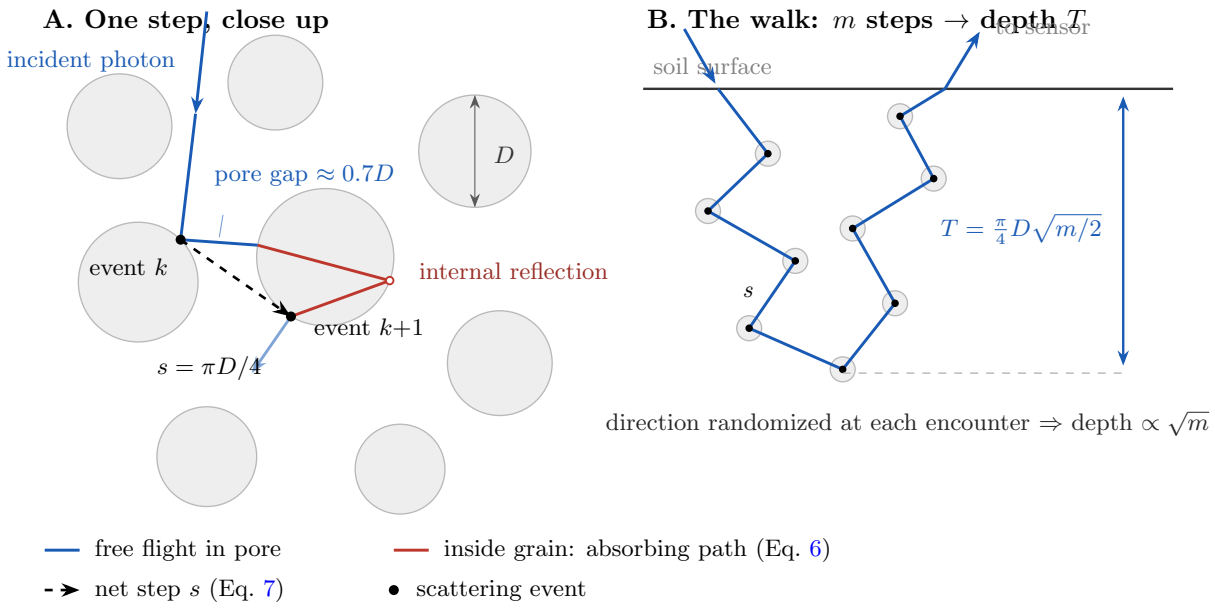


Figure 1: Physical meaning of the step length s in Eq. (7). Panel A: between scattering events k and $k{+}1$ the photon crosses a pore gap ($\approx 0.7D$ at typical porosity, blue), then refracts through the next grain, where the red in-grain segments constitute the absorbing path of Eq. (6) ($\approx 2D$ including internal reflections). The dashed arrow is the net event-to-event displacement s, the step of the random walk. Panel B: repeating this cycle m times with direction randomized at each event yields a random walk whose RMS vertical extent is the sampling depth T of Eq. (7).

The second term reads: entered the grain $(1 - S_e)$, survived one transit (Θ), exited $(1 - S_i)$, plus the geometric series of additional internal bounces hidden in the denominator. No individual photon's fate is decided; each branch contributes its probability-weighted share. Equations (6) and (7) apply the same philosophy to path length and displacement: $2D_i$ is the branch-averaged absorbing path per encounter (the external-reflection branch contributes zero absorbing path, and its ∼9% weight is part of why the constant is a convention), and $\pi D/4$ is the branch-averaged net displacement per event.

Deciding the branch for a particular event. Only a Monte Carlo ray tracer ever makes that decision, by sampling: draw the impact parameter (hence incidence angle θ) from the beam distribution, compute the polarization-averaged Fresnel reflectivity $R(\theta)$, draw a uniform random number ξ; if $\xi < R(\theta)$ the photon reflects externally in the specular direction off the local surface normal, otherwise it refracts in via Snell's law; at each internal wall hit, repeat with the internal Fresnel coefficient (unity beyond the critical angle); attenuate the photon weight by $e^{-\alpha\ell}$ along each internal segment. The analytic model of the ATBD replaces this entire decision tree with its expectation values and never branches at all.

Quantitative effect of the branching on s. The two branches give different step lengths and directions. An external bounce moves the photon only across the pore gap, $\sim 0.7D$, with broad angular redistribution off the convex surface. A through-transmission moves it pore gap plus grain chord, ~ 1.5 to $2D$, and is forward-biased: refraction through a sphere preferentially continues forward, which is why regolith grains have asymmetry parameter $g \sim 0.3$ to 0.6. The

branch-weighted mean displacement is then $\sim 0.1 \times 0.7D + 0.9 \times 1.7D \approx 1.6D$ per event, and the forward bias means direction is only partially randomized per event, so the effective isotropic step is larger still by the transport factor $1/(1-g)$. The adopted $\pi D/4 \approx 0.79D$ sits below this estimate, which indicates it should be read as a calibrated order-unity convention rather than a derived mean.

B.5 Caveats

Return to the surface. A detected photon is not a free random walker: it must return to the surface to be measured, so the correct problem is a first-passage walk in a half-space, and the depth distribution of returning walkers differs from free diffusion by another order-unity factor, likewise absorbed into the conventions.

Mineral-specific sampling depth. The step count in Eq. (7) is m_i, the encounters with mineral i alone, whereas the photon's true walk includes every grain it touches, overwhelmingly quartz and feldspar. T_i should therefore be read as a mineral-specific *sampling depth*, the effective thickness of the layer within which the detected photons accumulated their mineral-i absorption, rather than the literal physical penetration depth of the light. The algorithm applies the identical construction to quartz in Eq. (9), which keeps the bookkeeping consistent.

Insensitivity of the final product. As with the chord constant of Eq. (6), a uniform rescaling of the prefactor cancels in the final mass fractions: since $M \propto T \propto$ prefactor for every species including quartz, MR_i is invariant to the choice of $\pi/4$ and $1/2$ so long as they are applied uniformly. The constants matter only for the absolute areal masses, and, weakly, through the fact that m_i enters under a square root while the free parameter n enters Eq. (9) linearly, so the conventions and n are partially degenerate and are in effect calibrated together. Consistent with this, the retrieval is comparatively forgiving of errors in band depth, which enter T only as $\sqrt{\ln(1-BD)}$, but directly sensitive to the grain diameters, which enter linearly both here and in Eq. (9).

C Appendix: The areal-mass conversion of Eq. (8)

Equation (8),

$$M_i \;=\; 3\,T_i\,\rho_i \times 100^2 \quad [\mathrm{g\,m^{-2}}], \tag{39}$$

converts the sampling depth T_i of Eq. (7) into an areal mass. Three ingredients need justifying: the $\rho_i T_i$ product, the 100^2, and the factor 3. The first two are rigorous; the 3 is a convention with a physical reading. This appendix derives each, states the implicit assumptions, and cross-checks the construction against an exact alternative bookkeeping.

C.1 Elementary content: slab mass and units

The mass per unit surface area of a solid slab of thickness t and density ρ is $M = \rho\, t$. With t in cm and ρ in $\mathrm{g\,cm^{-3}}$, M comes out in $\mathrm{g\,cm^{-2}}$; multiplying by $(100\ \mathrm{cm/m})^2 = 100^2$ converts to $\mathrm{g\,m^{-2}}$. That accounts for everything in Eq. (8) except the 3: the equation asserts that the photons weighed an equivalent solid column of mineral i of thickness $3T_i$.

C.2 The factor 3: envelope versus RMS

T_i is the RMS of the vertical displacement of the walk, $T = \sqrt{\langle z^2 \rangle}$, the standard deviation of a zero-mean random variable. The RMS is not the extent of the region the photon ensemble actually

sampled: the depth distribution has a tail, and the layer containing essentially all of the sampled material extends to a few standard deviations. The factor 3 is the standard containment convention (for a Gaussian, $\pm 3\sigma$ captures 99.7% of samples), so the probed column is taken as $L = 3T_i$ and

$$M_i \;=\; \rho_i\, L \times 100^2 \;=\; 3\, T_i\, \rho_i \times 100^2. \tag{40}$$

The same 3 multiplies the quartz mass through Eq. (9), so, like every uniform constant in the chain, it cancels identically in MR_i; only the absolute areal masses carry it.

For the worked example of Section 8 (hematite-fine, $T = 1.6 \times 10^{-5}$ cm, $\rho = 5.3$ g cm^{-3}): $M = 3 \times 1.6 \times 10^{-5} \times 5.3 \times 10^4 \approx 2.5$ g m^{-2}. For scale, a solid hematite film of thickness T alone would weigh 0.85 g m^{-2}; the 3 is the envelope inflation.

C.3 The implicit assumption

Read literally, Eq. (8) treats the $3T_i$ column as *solid mineral i*: no porosity factor, no dilution by the other minerals sharing the layer. This is tolerable for two reasons. First, T_i was built from m_i alone, the mineral-i encounters, so it is already a mineral-specific equivalent depth that grows with that mineral's abundance; the dilution is encoded upstream. Second, the packing fraction and any other factor common to all species cancel in the delivered ratio MR_i. The absolute M_i in g m^{-2} should therefore not be over-interpreted; it is an intermediate bookkeeping quantity.

C.4 An independent cross-check: the straight-ray (Cauchy) bookkeeping

A fully rigorous alternative brackets the convention. Consider a straight vertical ray through a layer of thickness L containing grains of diameter D at number density N. The expected number of grains intersected is $m = N\sigma L$ with $\sigma = \pi D^2/4$, while the areal mass of those grains in the layer is $M = N\rho\,(\pi D^3/6)\,L$. Dividing,

$$\frac{M}{m} \;=\; \frac{\pi D^3/6}{\pi D^2/4}\,\rho \;=\; \frac{2}{3}\rho D \quad\Longrightarrow\quad M \;=\; \frac{2}{3}\,\rho\, D\, m \;=\; \frac{\rho\,\bar{d}}{3}, \tag{41}$$

where the last equality uses $m = \bar{d}/2D$ from Eq. (6). Each intersected grain contributes exactly one Cauchy mean chord ($2D/3$, Eq. (22)) of solid thickness, and the total mass is density times total in-grain path over 3. This is exact for a straight ray, and it scales *linearly* in m, whereas Eq. (8) scales as $\sqrt{m_i}$ because it routes through the diffusion depth. That difference is a genuine structural choice, not a cancelling constant.

Numerically, for the single-mineral worked example the two bookkeepings almost coincide in the delivered ratio. The straight-ray version gives $M_{\mathrm{hem}} = \rho\,\bar{d}/3 = 5.3 \times 1.65 \times 10^{-5}/3 = 2.9 \times 10^{-5}$ g cm^{-2} $= 0.29$ g m^{-2} and, with $n = 1$ and $m_q = \bar{m} = 0.082$, $M_q = \frac{2}{3}\,\rho_q D_q m_q = \frac{2}{3} \times 2.65 \times 0.01 \times 0.082 = 14.5$ g m^{-2}, hence $MR_{\mathrm{hem}} = 0.29/14.8 \approx 2.0\%$ versus the ATBD's 1.9%: with one detected mineral the m-dependence divides out of the ratio in both models. The two models diverge only when several minerals with different m_k are present, where the $\sqrt{m}$ weighting compresses the contrast between strong and weak absorbers relative to linear weighting.

C.5 Summary

$\rho T \times 100^2$ is elementary slab physics; the 3 is an envelope-containment convention that cancels in the delivered product; and the deeper choice hidden in this link, mass tied to the $\sqrt{m}$ diffusion depth rather than to the linear encounter count, is part of the semi-empirical calibration lineage of Clark et al. (2024), defensible in the weak-band regime and dominated in the error budget by grain-size uncertainty in any case.

D Appendix: A self-consistent alternative, inversion in single-scattering-albedo space

The closure inconsistency of Section 10 (item 0) can be removed without leaving the Hapke framework, following the prescription of Clark and Roush (1984): convert reflectance to single-scattering albedo *first*, and perform the mixture analysis in albedo space, where the intimate mixing rule is exactly linear. The nonlinearity is then handled where it belongs, in the $r \leftrightarrow w$ relation, and the free parameter n disappears. The algorithm is as follows.

Step 1: invert reflectance to albedo, channel by channel. Away from the opposition surge and with the isotropic-scattering approximation for the multiply scattered component, Eq. (2) reduces to

$$r(\mu_0, \mu) \;\approx\; \frac{w}{4\pi}\,\frac{\mu_0}{\mu_0 + \mu}\,H(\mu_0)\,H(\mu), \qquad H(\mu) \;\approx\; \frac{1 + 2\mu}{1 + 2\gamma\mu}, \qquad \gamma = \sqrt{1 - w}. \tag{42}$$

For each channel, given the measured reflectance and the solar and view geometry from L1B, Eq. (42) is a scalar equation for w, monotonic on $(0, 1)$, solved by bisection in microseconds. The output is the co-albedo spectrum $\alpha(\lambda) = 1 - w(\lambda)$. Surface roughness and porosity corrections (Hapke's shadowing function and porosity factor K) can be inserted here as optional refinements without altering the structure.

Step 2: continuum removal in co-albedo space. Fit a smooth continuum $\alpha_c(\lambda)$ over the shoulders of each diagnostic band and extract the band-center excess $\Delta\alpha_i = \alpha(\lambda_i) - \alpha_c(\lambda_i)$. This is the albedo-space analog of the band depth, but unlike BD it mixes linearly.

Step 3: linear unmixing. By the intimate mixing rule of Eq. (3), the mixture albedo is the cross-section-weighted average $w = \sum_k F_k w_k$ with $F_k = (M_k/\rho_k D_k)\big/\sum_j (M_j/\rho_j D_j)$. Since the bland matrix contributes no band-center excess, the observed excess at band i is

$$\Delta\alpha_i^{\mathrm{obs}} \;=\; F_i\,\Delta\alpha_i^{\mathrm{pure}} \quad\Longrightarrow\quad F_i \;=\; \frac{\Delta\alpha_i^{\mathrm{obs}}}{\Delta\alpha_i^{\mathrm{pure}}}, \tag{43}$$

where $\Delta\alpha_i^{\mathrm{pure}}$ is the pure-mineral co-albedo contrast, computable from the equivalent-slab model of Appendix A (from k_i and D_i via Θ, S_e, S_i) or, preferably, measured by inverting the library reflectance spectra themselves through Eq. (42). Minerals with multiple diagnostic bands yield an overdetermined linear system, solvable by weighted least squares, an immediate improvement over single-band depths.

Step 4: close the system and convert to mass. The matrix takes the remainder of the cross-section budget, $F_q = 1 - \sum_i F_i$. Inverting the definition of F_k gives mass fractions directly:

$$MR_i \;=\; \frac{F_i\,\rho_i D_i}{\sum_k F_k\,\rho_k D_k + F_q\,\rho_q D_q}, \tag{44}$$

with D_q from the Step 2 grain-size model exactly as in the current pipeline. Compare with Eqs. (9)–(10): the grain-count parameter n has disappeared, because the quartz term now enters through the same mixing rule as everything else, with no free bookkeeping.

Properties. Band saturation is handled naturally (w is bounded; the $r \leftrightarrow w$ map compresses correctly at strong absorption). Every constant in Eq. (44) is a physical parameter with laboratory provenance; the error budget becomes transparent. The remaining assumptions are inherited and explicit: per-class k_i, ρ_i, D_i (item 3 of Section 10), intimate mixture of equant grains (item 2; coatings remain outstanding), and the validity of Eq. (3) itself at high filling factor. Computationally the scheme is no heavier than the current chain.

E Appendix: A Monte Carlo benchmark, a compact research plan

The closure question of Section 10 (item 0) is decidable by a numerical experiment of modest scope. The objective is to measure, on synthetic soils with known ground truth, the band-depth-to-mass exponent actually realized by physically rigorous transport, and to quantify the retrieval error of the current chain against the albedo-space inversion of Appendix D.

Phase 1: forward model. Construct synthetic packed granular media by discrete-element sphere packing at filling factors 0.4 to 0.6: a polydisperse quartz/feldspar host (D_q = 50 to 500 μm) hosting absorber grains (hematite, goethite, kaolinite, calcite) at mass fractions from 0.1 to 20% and diameters spanning the Table 2 classes. Trace photons by geometric-optics Monte Carlo: Fresnel reflection and Snell refraction at every interface (polarization-averaged), Beer–Lambert attenuation along in-grain segments using measured complex refractive indices, and tallying of escaping photons by direction. This machinery resolves, by construction, every question raised in Appendices A and B: chord statistics, branching, path-length distributions, and return depths are outputs, not assumptions. Two morphology variants matter: dispersed absorber grains, and thin absorber *coatings* on host grains at equal mass loading, directly testing item 2 of Section 10. A cheaper cross-check forward model (DISORT or a two-stream code driven by Hapke-derived w and phase functions) should be run alongside to isolate dense-packing effects from closure effects.

Phase 2: retrieval intercomparison. Feed the synthetic spectra, convolved to EMIT channels and noise, through three retrievals: (a) the current L3 chain (Tetracorder-style band depths through Eqs. (5)–(10)); (b) the albedo-space inversion of Appendix D; (c) a full nonlinear Hapke fit as the ceiling. The decisive diagnostic is the exponent test: regress retrieved and true abundance against band depth, $MR \propto BD^{\beta}$, and locate β relative to the predictions of Table 2 (1/2, 1, 2) as a function of matrix albedo and absorber strength. Secondary metrics: bias and scatter of MR_i per mineral class, and cross-talk between coexisting minerals of unequal band depth.

Phase 3: laboratory anchor. Physical mixture series (sieved quartz plus hematite and goethite at controlled mass fractions and grain sizes; plus iron-coated versus acid-washed sand analogs), measured with a laboratory spectrometer under EMIT-like geometry. This ties the simulations to reality and revisits the calibration heritage of the group metadata with modern mixtures designed to span the band-depth dynamic range, which is where the exponent discrepancy expresses itself.

Phase 4: application and impact. Reprocess selected EMIT scenes over mineralogically characterized sites (gypsum at White Sands; Saharan and Middle Eastern sites with published soil mineralogy) with both retrievals, and propagate the differences to dust radiative forcing through an ESM mineral-speciation framework (e.g., Scanza et al., 2015; Li et al., 2020). The end-to-end sensitivity of the forcing to the retrieval closure is the number the community currently does not have.

Scope and risks. Phases 1 and 2 are a one-to-two year effort for a student or postdoc with strong radiative transfer skills; the Monte Carlo machinery has precedents in the planetary regolith literature and no exotic physics. The principal risk is that the geometric-optics forward model, like the retrievals, omits coherent dense-medium effects; the QCA literature (Tsang et al., 2000) bounds these, and at VSWIR size parameters they are expected to be small against the closure signal, which spans orders of magnitude in Table 2. The payoff is asymmetric: if the current chain's exponent is confirmed adequate within the calibrated range, the product gains a defense it presently lacks; if not, Appendix D is a drop-in replacement using the same inputs.

F Appendix: The matrix column of Eq. (9), reconstructed

This appendix expands Section 6.3: it derives the matrix encounter count from the encounter-share logic that the ATBD states in words, quantifies the three departures of the written Eq. (9) from that logic, and examines what the parameter n and the grain size D_q actually are.

F.1 Encounter shares

A photon random-walking through the mixed soil makes N_{tot} grain encounters in all. The probability that any given encounter involves species s is its share of the total geometric cross-section per unit volume, $N_s\sigma_s / \sum_{s'} N_{s'}\sigma_{s'}$ with $\sigma_s = \pi D_s^2/4$; this is the same weighting that underlies the Hapke mixing rule of Eq. (3). Hence, for the matrix (quartz/feldspar, q) relative to all detected minerals combined,

$$\frac{m_q}{\sum_k m_k} \;=\; \frac{N_q\,\sigma_q}{\sum_k N_k\,\sigma_k} \;\equiv\; \frac{1}{n_A}, \qquad n_A \;=\; \frac{\sum_k N_k D_k^2}{N_q D_q^2} \;=\; n\,\frac{\bar{D}^2_{\mathrm{min}}}{D_q^2}, \tag{45}$$

where $n = \sum_k N_k/N_q$ is the number ratio of the ATBD's definition and $\bar{D}^2_{\mathrm{min}} = \sum_k N_k D_k^2 / \sum_k N_k$. Two example soils show why the distinction between n and n_A matters. A clay-rich soil with 10% kaolinite by mass at 2 μm in 100 μm quartz has $n = \frac{0.10/(2.6\cdot 2^3)}{0.90/(2.65\cdot 100^3)} \approx 1.4\times 10^4$ but $n_A = n\,(2/100)^2 \approx 6$. A clean sand with 3% calcite at 20 μm has $n \approx 4$ and $n_A \approx 0.15$. The number ratio spans five orders of magnitude between these two soils; the cross-section ratio spans less than two, and its square root, which is what enters T_q, spans a factor of six. This is why a single calibrated constant can function at all, and it is a statement about n_A, not n.

F.2 The derived matrix column and the three departures

Applying Eq. (7) to the matrix encounters of Eq. (45):

$$T_q^{\mathrm{derived}} \;=\; \frac{\pi}{4}\,D_q\sqrt{\frac{m_q}{2}} \;=\; \frac{\pi}{4}\,D_q\sqrt{\frac{\sum_k m_k}{2\,n_A}}. \tag{46}$$

The ATBD's Eq. (9) instead reads $T_q^{\mathrm{ATBD}} = \frac{\pi}{4}\frac{D_q}{n}\sqrt{\sum_k m_k/2j}$. Their ratio is

$$\frac{T_q^{\mathrm{ATBD}}}{T_q^{\mathrm{derived}}} \;=\; \frac{\sqrt{n_A}}{n}\cdot\frac{1}{\sqrt{j}}, \tag{47}$$

which separates cleanly into a global constant and a pixel-dependent factor.

(i) Sum versus mean. The $1/\sqrt{j}$ is the mean-versus-sum choice. With the ATBD's mean, a pixel with j detected minerals has its matrix column, and hence M_q, reduced by $\sqrt{j}$ relative to the encounter-share derivation: factors of 1, 1.41, 1.73, and 2.0 for $j = 1$ to 4. Because M_q is typically the dominant term in the denominator of Eq. (10), every MR_i in a four-mineral pixel is inflated by up to a factor of 2 relative to a one-mineral pixel of identical mineralogy. This cannot be absorbed into calibration because j varies across the scene. The only rationale we can construct for the mean is that it prevents the detection of an additional weak mineral from deepening the matrix column; but under the ATBD's own definition of n as a ratio against *all* mineral grains, the matrix encounters are proportional to the *total* mineral encounters, and the sum is the consistent choice.

(ii) Placement of n. The factor $\sqrt{n_A}/n$ is a global constant. For a globally fixed parameter the placement of n inside or outside the root is therefore immaterial: the identity $(D_q/n)\sqrt{\bar{m}/2} = D_q\sqrt{\bar{m}/2n^2}$ shows that the ATBD's n is simply $\sqrt{n_A}$ under a relabeling, and calibration cannot distinguish them. The placement becomes a genuine error the moment n is assigned a physical value or made to vary with the Step 2 texture, which is the natural fix for item 1 of Section 10; at that point Eq. (46) with n_A from Eq. (45) is the form to implement.

(iii) The missing cross-section factor. Eq. (45) carries $D_q^2/\bar{D}_{\min}^2$ explicitly; the ATBD's form does not, and since D_q varies per pixel while n is fixed, the omission changes the scaling of the matrix mass with grain size. Under a fixed cross-section share (n_A fixed), Eq. (46) gives $M_q \propto \rho_q D_q$, linear, which is the ATBD's behavior and is also the scaling of the Hapke weights $M_q/(\rho_q D_q) \propto F_q$. Under a fixed *number* ratio (n fixed, the stated definition), $n_A \propto 1/D_q^2$ and Eq. (46) gives $M_q \propto \rho_q D_q^2$; the straight-ray mass route of Appendix C gives $\rho_q D_q^3$, the natural cubic scaling of mass with grain size at fixed grain count. The algorithm's linear scaling is therefore consistent only with reading n as a cross-section share. Since D_q is the one input whose uncertainty is propagated into the product, the exponent of its scaling law is not a detail: the delivered error bars assume the linear law.

F.3 What D_q is

Step 2 predicts the mass fractions of seven texture bins (clay, silt, and five sand classes from 1/16 to 2 mm) with a random forest trained on the ICRAF–ISRIC library. A cumulative distribution is assembled from the bin centers and D_q is its linearly interpolated 50th percentile: the median grain diameter of the bulk soil. Three consequences. It is the median of *all* particles, including the clay-sized fraction that is largely the detected minerals themselves, so in clay-rich pixels the "quartz/feldspar diameter" is set mostly by non-quartz grains; the label is a proxy that is clean only in sandy dust-source soils. It is the only pixel-varying grain size in the chain; every D_i is a class constant. And because M_q scales with it without a square root, it is the single most leveraged input in the product, which is why the ATBD builds its entire delivered uncertainty envelope from perturbing it by the Okin et al. (2025) two-sigma bounds of $-49.5/+80.0$ μm.

F.4 Summary

$\bar{m}$ is measured, D_q is estimated, and n is calibrated. The algebra connecting them departs from the ATBD's own encounter-share logic in the sum-versus-mean choice, which does not cancel and biases multi-mineral pixels by up to a factor of 2, and in the placement and meaning of n, which is harmless while n is a global constant and becomes wrong as soon as it is given physical content. The albedo-space formulation of Appendix D has no counterpart to any of these choices: the

matrix enters through $F_q = 1 - \sum_k F_k$ with no encounter proxy, no averaging over bands, and no exchange-rate parameter.

References

[1] Okin, G. S., Ochoa, F., Brodrick, P. G., Thompson, D. R., Keebler, A., Ehlmann, B. L., Mahowald, N., Miller, R. (2025). *EMIT L3 Algorithm: Aggregated Mineral Mass Fractions, Theoretical Basis*, Version 3.0, JPL D-104480.

[2] Hapke, B. (1981). Bidirectional reflectance spectroscopy: 1. Theory. *J. Geophys. Res.*, 86(B4), 3039–3054.

[3] Clark, R. N., and Roush, T. L. (1984). Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *J. Geophys. Res.*, 89, 6329–6340.

[4] Clark, R. N., et al. (2003). Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems. *J. Geophys. Res.*, 108(E12), 5131.

[5] Clark, R. N., et al. (2024). Imaging spectroscopy: Earth and planetary remote sensing with the PSI Tetracorder and expert systems from rovers to EMIT and beyond. *Planet. Sci. J.*, 5(12), 276. (Section 11.4.)

[6] Okin, G. S., Zhou, B., Brodrick, P. G., Ochoa, F., Thompson, D. R. (2025). Leveraging EMIT spectroscopy and machine learning to estimate soil composition and particle grain size in dust-producing regions. *Remote Sens. Environ.*, in review.

[7] Hapke, B. (1993). *Theory of Reflectance and Emittance Spectroscopy*. Cambridge University Press. (2nd ed. 2012.)

[8] Hapke, B. (2008). Bidirectional reflectance spectroscopy 6: Effects of porosity. *Icarus*, 195(2), 918–926.

[9] Tsang, L., Chen, C.-T., Chang, A. T. C., Guo, J., Ding, K.-H. (2000). Dense media radiative transfer theory based on quasicrystalline approximation with applications to passive microwave remote sensing of snow. *Radio Science*, 35(3), 731–749.

[10] Scanza, R. A., Mahowald, N., Ghan, S., Zender, C. S., Kok, J. F., Liu, X., et al. (2015). Modeling dust as component minerals in the Community Atmosphere Model. *Atmos. Chem. Phys.*, 15(1), 537–561.

[11] Li, L., Mahowald, N. M., Miller, R. L., et al. (2021). Quantifying the range of the dust direct radiative effect due to source mineralogy uncertainty. *Atmos. Chem. Phys.*, 21, 3973–4005.