When do supply risks add up? Dependence-aware aggregation of critical-mineral risks in alloys and multi-element systems

SUPPLEMENTARY INFORMATION

This Supplementary Information provides the technical derivations, calibration details, numerical diagnostics, and additional figures supporting the main-text framework. The main article is written for an interdisciplinary critical-mineral supply-risk readership and therefore retains only the methodological ingredients needed to understand the workflow and its implications for comparative screening. Here we document in detail the assumed link between elemental scores and latent incident probabilities, the Gaussian-copula pairwise approximation, the construction of the dependence matrix from companionality and country exposure, and the sensitivity analyses used in the worked example.

1. Technical derivation of the dependence-aware aggregation formula

We consider a single elemental supply-risk score and derive an alloy-level aggregation rule consistent with the system-level event that at least one required constituent becomes unavailable. The purpose of this section is not to claim that published supply-risk scores are calibrated probabilities, but to show how a transparent low-incident approximation can be used to preserve additive behavior as a baseline while quantifying dependence-induced corrections when needed.

The derivation proceeds in three steps. First, the elemental score is related to an unobserved incident probability through a monotone mapping, then approximated by a first-order local linearization in the low-incident regime. Second, alloy disruption is written as a union probability and approximated through pairwise inclusion–exclusion in the low-incident regime. Third, Gaussian-copula asymptotics provide a closed-form correction that interpolates between an additive regime and a dependence-dominated regime.

We consider a single elemental supply-risk indicator $R_i \equiv R(E_i) \geq 0$. The event A_i denotes the supply-disruption event associated with element E_i, and

$$p_i = \mathbb{P}(A_i)$$

denotes its unobserved probability over a reference horizon. The published score R_i is not assumed to be a calibrated probability. We assume only that p_i is related to R_i by an element-independent monotone function f:

$$p_i = f(R_i)$$

Because the derivation is restricted to a low-incident regime, we use the first-order local approximation

$$p_i \approx \varepsilon R_i$$

where ε is a global scale parameter common to all elements. This approximation is not a calibration of R_i into an absolute probability; it is a modelling convention that allows the aggregation of scores to obey the probability laws expected for the underlying incident events.

For an alloy $\mathcal{A} \equiv \{E_{i_1}, \cdots, E_{i_a}\}$, the alloy-level probability is defined as the probability that at least one constituent element is hit:

$$P(\mathcal{A}) \equiv \mathbb{P}\left(\bigcup_{i=1}^{N} A_i\right)$$

Because incidents across elements are not independent, we model dependence using a Gaussian copula: we introduce a latent Gaussian vector $\boldsymbol{Z} \equiv (Z_i)_{i\in\{i_1,\ldots,i_a\}} \sim \mathcal{N}(\mathbf{0}, \Sigma_E)$ whose correlation matrix Σ_E is constructed in Section 2. Each marginal probability $\mathbb{P}(A_{i_k})$ is mapped to a threshold

$$q_{i_k} = \bar{\Phi}^{-1}\left(\mathbb{P}(A_{i_k})\right)$$

so that

$$\mathbb{P}(Z_{i_k} > q_{i_k}) = \mathbb{P}(A_{i_k})$$

The multivariate normal cumulative distribution function then yields $P(\mathcal{A}) = 1 - \mathbb{P}(Z_{i_1} < q_{i_1}, \ldots, Z_{i_a} < q_{i_a})$, evaluated numerically when needed by Monte Carlo / Cholesky sampling.

1.1. Gaussian-copula dependence between elements (latent Gaussian vector).

The previous subsection defined the latent incident probabilities p_i associated with elemental scores $R_i \equiv R(E_i) \geq 0$. We now introduce the Gaussian-copula representation used to combine these marginals with an element–element dependence matrix. This representation does not assume that the original scores are Gaussian. It only assumes that the joint occurrence of the latent incident events can be represented through threshold exceedances of a correlated Gaussian vector.

One defines the transformation corresponding to the inverse of the distribution function of the normal distribution $\Phi: x \mapsto (2\pi)^{-\frac{1}{2}} \int_{z=-\infty}^{x} \varphi(z)dz$ with $\varphi: z \mapsto e^{-\frac{z^2}{2}}$ and $\bar{\Phi}: x \mapsto 1 - \Phi(x) = (2\pi)^{-\frac{1}{2}} \int_{z=x}^{+\infty} \varphi(z)dz$:

$$\bar{\Phi}^{-1}: \begin{cases} [0,1] \to [-\infty, +\infty] \\ p \mapsto q \text{ such as } P(z > q) = p \Leftrightarrow \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{z=q}^{+\infty} e^{-\frac{z^2}{2}} dz = p \end{cases} \tag{1-1}$$

Let now $\boldsymbol{q} \in \mathbb{R}^n$ be defined by

$$q_i \equiv \bar{\Phi}^{-1}(p_i) = \bar{\Phi}^{-1}(\varepsilon R(E_i)) \tag{1-2}$$

So that $\mathbb{P}(Z_i > q_i) = \mathbb{P}(A_i)$ for a standard normal Z_i.

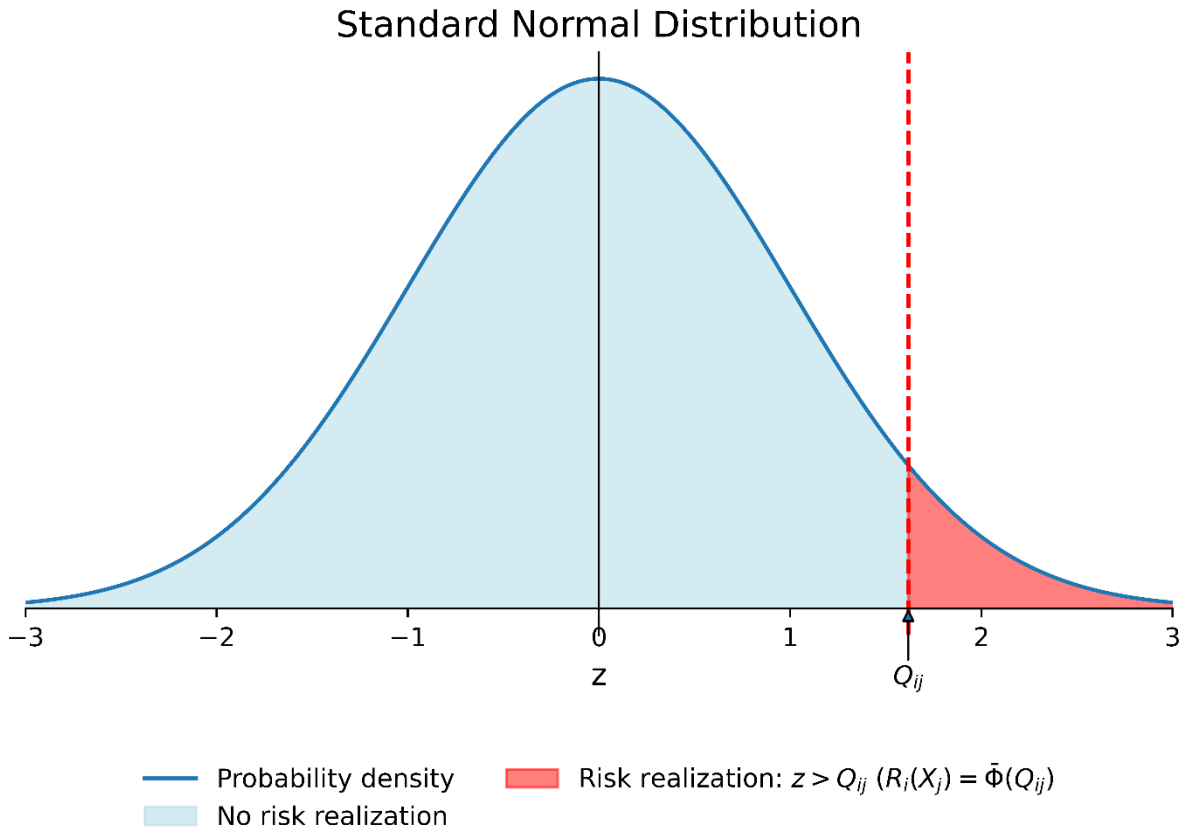

Figure 1: Latent Gaussian threshold associated with an incident probability $p_i = \bar{\Phi}(q_i)$

The Gaussian curve represents the latent standard normal variable Z_i. For a given incident probability p_i, the threshold $q_i = \bar{\Phi}^{-1}(p_i)$ is chosen so that the upper tail $\mathbb{P}(Z_i > q_i)$ equals p_i. The threshold construction is used only inside the Gaussian-copula model; it does not imply that the original supply-risk score is Gaussian.

Through the transformation $q_i = \bar{\Phi}^{-1}(p_i)$, latent incident probabilities are transformed into normal quantiles. This allows the joint occurrence of incident events to be represented as threshold exceedances of a correlated Gaussian vector. Dependence between elements is encoded in the element–element correlation matrix $\rho_{ij} \equiv (\Sigma_E)_{ij}$ of the latent Gaussian vector $\boldsymbol{Z} \sim \mathcal{N}(\mathbf{0}, \Sigma_E)$. In this work, Σ_E is constructed from joint production and companionality (Section 2, see Eq. (2-15)).

To illustrate the practical effect of correlations between elements, **Figure 2** shows a simple two-dimensional example. Monte Carlo samples are drawn of a bivariate normal vector (z_1, z_2) with prescribed quantiles Q_1 and Q_2, first in the independent case and then with a positive correlation encoded in Σ_E. The red region corresponds to realizations where at least one of the two risks materializes; comparing the two panels highlights how positive correlation concentrates events and reduces the overall probability of "at least one" risk, so that neglecting correlations would overestimate the aggregated risk.

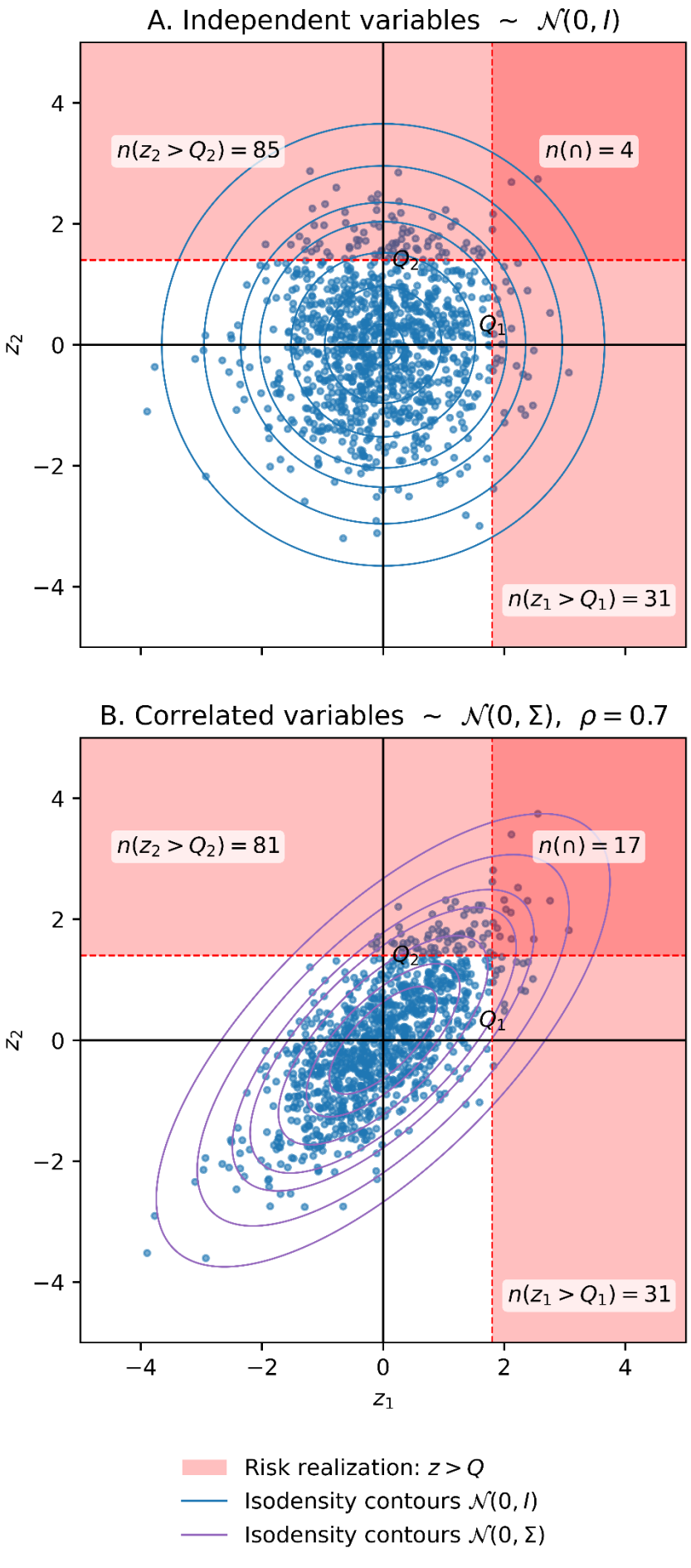
A. Independent variables ~ $\mathcal{N}(0, I)$
$n(z_2 > Q_2) = 85$
$n(\cap) = 4$
$n(z_1 > Q_1) = 31$
Q_2
Q_1
B. Correlated variables ~ $\mathcal{N}(0, \Sigma)$, $\rho = 0.7$
$n(z_2 > Q_2) = 81$
$n(\cap) = 17$
$n(z_1 > Q_1) = 31$
Q_2
Q_1
z_1
z_2
Risk realization: $z > Q$
Isodensity contours $\mathcal{N}(0, I)$
Isodensity contours $\mathcal{N}(0, \Sigma)$

Figure 2: Simulated latent Gaussian events with and without positive dependence

The risk R_1 (resp. R_2) is realized whenever $z_1 > Q_1$ (resp. $z_2 > Q_2$). Only the dots in the red area correspond to the risk realization. The contour lines correspond to equiprobability of one (at least) risk realization. The quantity n stands for the number of dots satisfying the condition $z_i > Q_i$, $i = 1$ or $i = 2$. The number of dots satisfying both conditions at a time is denoted by $n(\cap)$.

A. The latent variables associated with E_i and E_j are independent.

B. The latent variables associated with E_i and E_j are correlated with the respect of $\mathcal{N}(\mathbf{0}, \Sigma_E)$ – here $\Sigma_E = \begin{pmatrix} 1 & 0.7 \\ 0.7 & 1 \end{pmatrix}$. The correlation affects the probability of the global risk – by having both risks realized in at the same event, instead of two different events, decreasing the number of dots in the red area (visible as the increase of $n(\cap)$). Neglecting the correlation between elements would lead to overestimate the global risk.

The Gaussian random variable that is defined has the following probability density:

$$F: \begin{cases} \mathbb{R}^a \to \mathbb{R} \\ \boldsymbol{Z} \mapsto \dfrac{1}{(2\pi)^{\frac{a}{2}} \det(\Sigma_E)^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \boldsymbol{Z}^\top {\Sigma_E}^{-1} \mathbf{Z}\right) \end{cases} \tag{1-3}$$

This allows to express the probability that no incident affects any element in the alloy $\mathcal{A} \equiv \{E_{i_1}, \cdots, E_{i_a}\}$, based on the thresholds q_{i_k} defined in Eq. (1-2):

$$\begin{aligned} 1 - P(\mathcal{A}) &\equiv \mathbb{P}\left(Z_{i_1} < q_{i_1}, \dots, Z_{i_a} < q_{i_a}\right) \\ &= \int_{-\infty}^{q_{i_1}} \dots \int_{-\infty}^{q_{i_a}} \frac{1}{(2\pi)^{\frac{a}{2}} \det(\Sigma_E)^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \boldsymbol{Z}^\top {\Sigma_E}^{-1} \mathbf{Z}\right) d\boldsymbol{Z} \end{aligned} \tag{1-4}$$

This integral cannot be directly calculated, but it can be approximated by the Monte Carlo method (by using the Cholesky method to simulate the multivariate random variables). Indeed, given that Σ_E is a symmetric positive definite matrix, it is known that it can be decomposed in the form $\Sigma_E = LL^\top$ with L being a triangular matrix. Moreover, one of the properties of a multivariate normal distribution is that if $A \in \mathcal{M}_{a,a}(\mathbb{R}), \boldsymbol{Y} \in \mathbb{R}^a, \boldsymbol{Y} \sim \mathcal{N}(\boldsymbol{\mu}, C)$, then $A\boldsymbol{Y} \sim \mathcal{N}(A\boldsymbol{\mu}, ACA^\top)$.

By generating a standard Gaussian random vector $\boldsymbol{U} \in \mathbb{R}^q$, the vector $L\boldsymbol{U}$ thus follows a distribution $\mathcal{N}(\mathbf{0}, LL^\top) = \mathcal{N}(\mathbf{0}, \Sigma_E)$.

To obtain an approximate numerical value of $P(\mathcal{A})$, and hence $R(\mathcal{A}) \equiv P(\mathcal{A})/\varepsilon$, once the matrix L is calculated, it is possible to quickly generate many Gaussian vectors $\boldsymbol{Z}$ with covariance matrix Σ_E, then calculate the proportion of them contained in the hypervolume $Z_{i_1} < q_{i_1}, \dots, Z_{i_a} < q_{i_a}$. This allows for deducing the probability $\mathbb{P}\left(Z_{i_1} < q_{i_1}, \dots, Z_{i_a} < q_{i_a}\right)$, then $R(\mathcal{A})$.

The same latent-vector representation is then used below for the alloy-level aggregation problem, where several elemental incident events must be combined under dependence.

1.2. From elemental scores to latent probabilities and back to a comparative indicator

This subsection addresses the following question: given a vector of elemental scores R_i, can one define an alloy-level indicator $R_{\text{alloy}}(\mathcal{A})$ that remains comparable with the elemental values while respecting the probability laws governing the underlying disruption events?

Many published risk values are scores, not calibrated probabilities. We therefore distinguish three objects: the published elemental score R_i, the unobserved incident probability $p_i = \mathbb{P}(A_i)$, and the extended alloy-level score $R_{\text{alloy}}(\mathcal{A})$. The connection between R_i and p_i is not known exactly. We assume that

$$p_i = f(R_i)$$

where f is element-independent and monotone. If f is regular and the relevant probabilities are small, its first-order local approximation gives:

$$p_i \approx \varepsilon R_i$$

This approximation preserves relative comparisons on the original score scale while allowing the aggregation step to be performed in probability space. After computing the alloy-level probability p_{alloy}, dividing by the same ε rescales the result back to the original indicator scale:

$$R_{\text{alloy}}(\mathcal{A}) \equiv \frac{p_{\text{alloy}}}{\varepsilon}$$

The resulting $R_{\text{alloy}}(\mathcal{A})$ is therefore not a calibrated probability forecast. It is a comparative indicator designed to rank elements and multi-element systems on a common score-consistent scale.

The practical procedure is therefore: infer latent probabilities up to a global scale factor, apply the aggregation in probability space, and rescale the resulting alloy-level probability back to the original score scale.

Intuitively, there are two opposite "modes" in the combination of the R_i depending on the correlation of the elements:

- If A and B are independent (correlation is zero), then $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A)\mathbb{P}(B) \approx \mathbb{P}(A) + \mathbb{P}(B)$ if $\mathbb{P}(A)$ and $\mathbb{P}(B)$ are small (see **Figure 2**: the narrower each red band, the more negligible the ratio represented by the area of their intersection (dark red)). Low correlation means a "sum mode".
- If A and B are absolutely correlated, then $A \subset B$ (or $B \subset A$), $\mathbb{P}(A \cup B) = \mathbb{P}(A) = \max\big(\mathbb{P}(A), \mathbb{P}(B)\big)$ (to include the case $B \subset A$). High correlation means a "max mode".

The next step is to identify an analytical function that captures both aggregation modes, along with the threshold separating them. This provides a coarse yet practical approximation for combining risks, enabling a decision between additive (sum-like) and dominant-factor (max-like) behavior.

To be able to perform a first order development for small probabilities (meaning large x), an asymptotic development of Gaussian integrals is needed. Since $\varphi'(t) = -t\varphi(t)$, an integration by part, for $x > 0$, gives:

$$\int_{t=x}^{+\infty} \varphi(t)dt = \int_{t=x}^{+\infty} \frac{-1}{t}\varphi'^{(t)}dt = \left[\frac{-1}{t}\varphi(t)\right]_{t=x}^{+\infty} - \int_{t=x}^{+\infty} \frac{1}{t^2}\varphi(t)dt = \frac{\varphi(x)}{x} - \int_{t=x}^{+\infty} \frac{\varphi(t)}{t^2}dt$$

$$1 - \Phi(x) = \frac{\varphi(x)}{x} - \int_{t=x}^{+\infty} \frac{\varphi(t)}{t^2}dt \tag{1-5}$$

Obviously in Eq. (1-5), $\int_{t=x}^{+\infty} \frac{\varphi(t)}{t^2}dt \geq 0$ and therefore:

$$1 - \Phi(x) \leq \frac{\varphi(x)}{x} \tag{1-6}$$

In $\int_{t=x}^{+\infty} \frac{\varphi(t)}{t^2}dt$, $t \geq x$, so:

$$\int_{t=x}^{+\infty} \frac{\varphi(t)}{t^2}dt \leq \frac{1}{x^2}\int_{t=x}^{+\infty} \varphi(t)dt = \frac{1}{x^2}\big(1 - \Phi(x)\big)$$

Using it in Eq. (1-5):

$$1 - \Phi(x) \geq \frac{\varphi(x)}{x} - \frac{1}{x^2}\big(1 - \Phi(x)\big)$$

So:

$$\big(1 - \Phi(x)\big)\left(1 + \frac{1}{x^2}\right) \geq \frac{\varphi(x)}{x}$$

and since $1 + \frac{1}{x^2} \geq 0$:

$$1 - \Phi(x) \geq \frac{\varphi(x)}{x + \frac{1}{x}} \tag{1-7}$$

This allows to frame the Gaussian tail:

$$\forall x > 0, \frac{1}{1 + \frac{1}{x^2}}\frac{\varphi(x)}{x} \leq 1 - \Phi(x) \leq \frac{\varphi(x)}{x} \tag{1-8}$$

This result is known as “Mill’s inequality”.

Let q_i be such as $p_i = \varepsilon r_i = \frac{1}{\sqrt{2\pi}}\int_{t=q_i}^{+\infty} e^{-\frac{t^2}{2}}dt = 1 - \Phi(q_i)$, as in Eq. (1-1) and (1-2). If p_i is small, q_i is big enough to approximate $\frac{1}{1+\frac{1}{{q_i}^2}} \approx 1$ and then:

$$p_i \approx \frac{\varphi(q_i)}{q_i} = \frac{e^{-\frac{{q_i}^2}{2}}}{\sqrt{2\pi} q_i} \tag{1-9}$$

To calculate these integrals, one proceeds by "bands", fixing one parameter. This approach relies on the fact that for a couple of Gaussian variables, any "conditional slice" remains Gaussian.

Let now denote the events $A_i = \{Z_i > q_i\}$ and $\overline{\Phi} \equiv 1 - \Phi$:

$$\mathbb{P}(A_i \cap A_j) = \iint\limits_{\substack{v>q_j \\ u>q_i}} \frac{1}{(2\pi)^{\frac{2}{2}} \det(\Sigma)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\begin{pmatrix} u \\ v \end{pmatrix}^{\top} \Sigma^{-1} \begin{pmatrix} u \\ v \end{pmatrix}\right) dvdu \tag{1-10}$$

With $\Sigma = \begin{pmatrix} 1 & \rho_{ij} \\ \rho_{ij} & 1 \end{pmatrix}$ and therefore $\det(\Sigma) = 1 - {\rho_{ij}}^2$ and $\Sigma^{-1} = \frac{1}{1-{\rho_{ij}}^2}\begin{pmatrix} 1 & -\rho_{ij} \\ -\rho_{ij} & 1 \end{pmatrix}$.

$$\mathbb{P}(A_i \cap A_j) = \frac{1}{2\pi\sqrt{1-{\rho_{ij}}^2}} \iint\limits_{\substack{v>q_j \\ u>q_i}} \exp\left(-\frac{1}{2}\left(\frac{u^2 - 2\rho_{ij}uv + v^2}{1-{\rho_{ij}}^2}\right)\right) dvdu \tag{1-11}$$

The joint law of $\{Z_i, Z_j\}$ is therefore $f_{i,j}(u,v) \equiv \frac{1}{2\pi\sqrt{1-{\rho_{ij}}^2}} \exp\left(-\frac{1}{2}\left(\frac{u^2-2\rho_{ij}uv+v^2}{1-{\rho_{ij}}^2}\right)\right)$.

This joint law can be factorized with the conditional $f_{Z_j}(v) = \varphi(v)$:

$$f_{Z_i|Z_j=v}(u) = \frac{f_{i,j}(u,v)}{f_{Z_j}(v)} = \frac{1}{\sqrt{2\pi\left(1-{\rho_{ij}}^2\right)}} \exp\left(-\frac{1}{2}\left(\frac{u^2 - 2\rho_{ij}uv + v^2}{1-{\rho_{ij}}^2} - v^2\right)\right)$$

$$f_{Z_i|Z_j=v}(u) = \frac{1}{\sqrt{2\pi\left(1-{\rho_{ij}}^2\right)}} \exp\left(-\frac{1}{2}\frac{\left(u-\rho_{ij}v\right)^2}{1-{\rho_{ij}}^2}\right) \sim \mathcal{N}\left(\rho_{ij}v, 1-{\rho_{ij}}^2\right) \tag{1-12}$$

As expected, the "conditional slice" that appears in Eq. (1-12) follows a Gaussian law.

$$\mathbb{P}(A_i \cap A_j) = \int\limits_{v=q_j}^{+\infty} \left[f_{Z_i|Z_j=v}(u)\right] f_{Z_j}(v)dv = \int\limits_{v=q_j}^{+\infty} \mathbb{P}\left(Z_i > q_i \middle| Z_j = v\right)\varphi(v)dv \tag{1-13}$$

For a normal law, $\mathbb{P}(X > q) = \overline{\Phi}\left(\frac{q-\mu}{\sigma}\right)$, with here $\mu = \rho_{ij}v$ and $\sigma = 1 - {\rho_{ij}}^2$.

So:

$$\mathbb{P}\left(Z_i > q_i \middle| Z_j = v\right) = \overline{\Phi}\left(\frac{q_i - \rho_{ij}v}{\sqrt{1-{\rho_{ij}}^2}}\right) \tag{1-14}$$

And then:

$$\mathbb{P}(A_i \cap A_j) = \int_{v=q_j}^{+\infty} \bar{\Phi}\left(\frac{q_i - \rho_{ij} v}{\sqrt{1-{\rho_{ij}}^2}}\right) \varphi(v) dv \tag{1-15}$$

Let $m_{ij} = \min(q_i, q_j)$, $M_{ij} = \max(q_i, q_j)$. Having the possibility to switch i and j in the Eq. (1-15), it comes:

$$\boxed{\mathbb{P}(A_i \cap A_j) = \int_{v=m_{ij}}^{+\infty} \bar{\Phi}\left(\frac{M_{ij} - \rho_{ij} v}{\sqrt{1-{\rho_{ij}}^2}}\right) \varphi(v) dv} \tag{1-16}$$

This can be developed in a double integral, and if the approximation (1-9) for $\bar{\Phi}$ is used, this will be approximated by a simple integral that includes a product $\varphi(x)\varphi(y) \propto e^{x^2+y^2}$.

$$\begin{aligned}\left(\frac{M_{ij} - \rho_{ij} v}{\sqrt{1-{\rho_{ij}}^2}}\right)^2 + v^2 &= \frac{1}{1-{\rho_{ij}}^2}\left({M_{ij}}^2 - 2M_{ij}\rho_{ij} v + \left(\rho_{ij} v\right)^2 + \left(1-{\rho_{ij}}^2\right)v^2\right) \\ &= \frac{1}{1-{\rho_{ij}}^2}\left({M_{ij}}^2 - 2M_{ij}\rho_{ij} v + v^2\right) \\ &= \frac{1}{1-{\rho_{ij}}^2}\left[\left(v - M_{ij}\rho_{ij}\right)^2 + {M_{ij}}^2\left(1-{\rho_{ij}}^2\right)\right]\end{aligned} \tag{1-17}$$

Let $t \equiv \frac{v - M_{ij}\rho_{ij}}{\sqrt{1-{\rho_{ij}}^2}}$:

$$\boxed{\left(\frac{M_{ij} - \rho_{ij} v}{\sqrt{1-{\rho_{ij}}^2}}\right)^2 + v^2 = t^2 + {M_{ij}}^2} \tag{1-18}$$

We now apply the change of variable from v to t:

$$dv = \sqrt{1-{\rho_{ij}}^2}\, dt$$

The minimum of t, when $v = m_{ij}$, is $\delta_{ij} = \frac{m_{ij} - M_{ij}\rho_{ij}}{\sqrt{1-{\rho_{ij}}^2}}$

$$\begin{aligned}\mathbb{P}(A_i \cap A_j) = \int_{t=\delta_{ij}}^{+\infty} \bar{\Phi}\left(\sqrt{1-{\rho_{ij}}^2} M_{ij} - \rho_{ij} t\right) \varphi\left(t\sqrt{1-{\rho_{ij}}^2}\right. \\ \left. + M_{ij}\rho_{ij}\right)\sqrt{1-{\rho_{ij}}^2}\, dt\end{aligned} \tag{1-19}$$

When v varies in $[m_{ij}, +\infty[$, $M_{ij} - \rho_{ij} v$ varies in $]-\infty, M_{ij} - \rho_{ij} m_{ij}]$ that includes 0. It is reached at t_0 such that $\sqrt{1-{\rho_{ij}}^2} M_{ij} - \rho_{ij} t_0 = 0$.

For simplification purposes, let's denote:

- $\rho \equiv \rho_{ij}$
- $M \equiv M_{ij}$

- $\delta \equiv \delta_{ij}$
- $a(t) \equiv \rho M + \sqrt{1-\rho^2}t$ (also noted a when t does not matter)
- $b(t) \equiv \sqrt{1-\rho^2}M - \rho t$ (also noted b when t does not matter);
    - since M is large, b is also large; $\frac{1}{b} = \frac{1}{\sqrt{1-\rho^2}M}\frac{1}{1-\frac{\rho t}{\sqrt{1-\rho^2}M}} = \frac{1}{\sqrt{1-\rho^2}M}\big(1+O(1)\big)$.
    - $b(t_0) = 0$.
- Eq. (1-18), combined with the change of integration variable made in (1-19) prove that $a(t)^2 + b(t)^2 = M^2 + t^2$ and therefore that $\varphi(a)\varphi(b) = \varphi(M)\varphi(t)$
- $J \equiv \sqrt{1-\rho^2}\int_{t=\delta}^{+\infty} \bar{\Phi}\big(b(t)\big)\,\varphi\big(a(t)\big)dt$
- $r_{min}(\mathrm{i},\mathrm{j}) \equiv \min\{r_i, r_j\}$ and $p_{min}(\mathrm{I},\mathrm{j}) \equiv \min\{p_i, p_j\}$, and thus $\frac{\varphi(M)}{M} \approx p_{min}$

The first step is to split the integral J on $[\delta, t_0]$ and $[t_0, +\infty[$. Since:

$$a(t_0) = \rho M + \sqrt{1-\rho^2}\frac{\sqrt{1-\rho^2}M}{\rho} = \frac{M}{\rho}$$

the second integral is dominated, using $\bar{\Phi}(b) \le 1$, inequality (1-8) on $\bar{\Phi}(M/\rho)$, approximation (1-9) on $p_m \approx \varphi(M)/M$:

$$\begin{aligned} 0 \le \sqrt{1-\rho^2}\int_{t=t_0}^{+\infty} \varphi\big(a(t)\big)\,\bar{\Phi}\big(b(t)\big)dt &\le \sqrt{1-\rho^2}\int_{u=a(t_0)}^{+\infty} \varphi(u)\frac{du}{\sqrt{1-\rho^2}} \\ &= \bar{\Phi}\left(\frac{M}{\rho}\right) \le \frac{\varphi\left(\frac{M}{\rho}\right)}{M/\rho} \approx \rho\,\frac{p_m}{e^{-\frac{M^2}{2}}}e^{-\frac{M^2}{(2\rho^2)}} \\ &= \rho p_m \exp\left(-\frac{M^2}{2}\left(\frac{1}{\rho^2}-1\right)\right) = o(p_m) \end{aligned} \tag{1-20}$$

On the interval $[\delta, t_0[$, $b(t) > 0$ so inequality (1-8) is valid:

$$\frac{1}{1+\frac{1}{b^2}}\frac{\varphi(b)}{b} = \frac{\varphi(b)}{b}\left(1+O\left(\frac{1}{M^2}\right)\right) \le \bar{\Phi}(b) \le \frac{\varphi(b)}{b} \tag{1-21}$$

It can be deduced:

$$\bar{\Phi}(b) = \frac{\varphi(b)}{b} + O\left(\frac{\varphi(b)}{M^2}\right) \tag{1-22}$$

Using $\varphi(a)\varphi(b) = \varphi(M)\varphi(t)$:

$$\int_{t=\delta}^{t_0} O\left(\frac{\varphi(b)}{M^2}\right)\varphi\big(a(t)\big)dt = \int_{t=\delta}^{t_0} O\left(\frac{\varphi(M)}{M^2}\right)\varphi(t)dt = O\left(\frac{\varphi(M)}{M^2}\right) \tag{1-23}$$

It is now possible to assess J from (1-20)(1-23) and (1-23):

$$J = o\left(\frac{\varphi(M)}{M}\right) + \sqrt{1-\rho^2}\int_{t=\delta}^{t_0}\frac{\varphi\big(b(t)\big)}{b(t)}\varphi\big(a(t)\big)dt + O\left(\frac{\varphi(M)}{M^2}\right)$$

$$= \sqrt{1-\rho^2}\int_{t=\delta}^{t_0}\frac{\varphi(M)}{b(t)}\varphi(t)dt + o\left(\frac{\varphi(M)}{M}\right)$$

$$= \sqrt{1-\rho^2}\int_{t=\delta}^{t_0}\frac{1}{\sqrt{1-\rho^2}}\left(1+\frac{1}{1-\frac{\rho t}{\sqrt{1-\rho^2}M}}\right)\frac{\varphi(M)}{M}\varphi(t)dt + o\left(\frac{\varphi(M)}{M}\right)$$

$$= \frac{\varphi(M)}{M}\int_{t=\delta}^{t_0}\left(1+\frac{1}{1-\frac{\rho t}{\sqrt{1-\rho^2}M}}\right)\varphi(t)dt + o\left(\frac{\varphi(M)}{M}\right)$$

$$J = \frac{\varphi(M)}{M}\left(\int_{t=\delta}^{t_0}\varphi(t)dt + O\left(\frac{1}{M}\right)\right) + o\left(\frac{\varphi(M)}{M}\right) \qquad (1\text{-}24)$$

Since $t_0 = \frac{\sqrt{1-\rho^2}M}{\rho}$ is increasing to infinity with M, $\int_{t=\delta}^{t_0}\varphi(t)dt = \int_{t=\delta}^{+\infty}\varphi(t)dt + O(\varphi(t_0)/t_0) = \bar{\Phi}(\delta) + O(1/M)$.

From $p_k \approx \frac{\varphi(q_k)}{q_k}$, it comes in particular $p_m \approx \frac{\varphi(M)}{M}$; as a conclusion:

$$\boxed{J = \frac{\varphi(M)}{M}\big(\bar{\Phi}(\delta) + o(1)\big) = p_m\big(\bar{\Phi}(\delta) + o(1)\big) \xrightarrow{\varepsilon\to 0} p_m\bar{\Phi}(\delta)} \qquad (1\text{-}25)$$

We now return to the alloy-level aggregation problem. For small probabilities, it is reasonable to neglect intersections of three or more sets that would be $o(\varepsilon^2)$ (see **Figure 2**):

$$\begin{aligned}\frac{1}{\varepsilon}\mathbb{P}\left(\bigcup_k A_k\right) &\approx \frac{1}{\varepsilon}\sum_k \mathbb{P}(A_k) - \frac{1}{\varepsilon}\sum_{i<j}\mathbb{P}\left(A_i \cap A_j\right)\\ &= \sum_k r_k - \sum_{i<j}\bar{\Phi}\big(\delta_{ij}\big)\cdot r_{min}(\mathrm{i},\mathrm{j})\end{aligned} \qquad (1\text{-}26)$$

This expression identifies the "trigger" term, $\bar{\Phi}(\delta)$, responsible for the switch between two modes, sum or max. The switch takes place when $\delta = 0$. As a reminder, $\delta = \frac{m-M\rho}{\sqrt{1-\rho^2}}$. This expression contains m and M, which are very inconvenient to handle, since they are only implicitly defined by $p = \bar{\Phi}(q)$, with p_i containing the unknown parameter ε. It is therefore necessary to get rid of the q.

Inequality (1-8) on $q > 0$ is:

$$\frac{\varphi(q)}{q+\frac{1}{q}} \leq p = \bar{\Phi}(q) \leq \frac{\varphi(q)}{q}$$

This inequality implies:

$$\log\left(\frac{q}{\varphi(q)}\right) = \log(q) + \frac{q^2}{2} + \frac{1}{2}\log(2\pi) \leq \log\left(\frac{1}{p}\right) \leq \log\left(\frac{q+\frac{1}{q}}{\varphi(q)}\right) \qquad (1\text{-}27)$$
$$= \log\left(q+\frac{1}{q}\right) + \frac{q^2}{2} + \frac{1}{2}\log(2\pi)$$

From $\frac{q^2}{2} \leq \log\left(\frac{1}{p}\right) - \log(q) - \frac{1}{2}\log(2\pi) \leq \log\left(\frac{1}{p}\right)$, it comes:

$$q \leq \sqrt{2\log\left(\frac{1}{p}\right)} \qquad (1\text{-}28)$$

Intuitively, the bigger q, the more negligible is $\log\left(q+\frac{1}{q}\right)$ compared to $\frac{q^2}{2}$. The second inequality should therefore make it possible to prove an equivalence when $q \to +\infty$. Let us assume there exists $\zeta > 0$ such that $\forall \eta > 0, \exists q \geq \eta, q \leq \sqrt{2\log\left(\frac{1}{p}\right)}(1-\zeta)$.

Then $\frac{q^2}{2} \leq \log\left(\frac{1}{p}\right)(1-\zeta)^2$ and injecting this into (1-27), $\log\left(\frac{1}{p}\right) \leq \log\left(\frac{1}{p}\right)(1-\zeta)^2 + \log\left(q+\frac{1}{q}\right) + \frac{1}{2}\log(2\pi) \leq \log\left(\frac{1}{p}\right)(1-\zeta)^2 + \frac{1}{2}\log\left(8\log\left(\frac{1}{p}\right)\right) + \frac{1}{2}\log(2\pi)$ which is in contradiction with $\lim_{q\to+\infty} \log\left(\frac{1}{p}\right) = +\infty$.

This proves the following result:

$$q \sim_{+\infty} \sqrt{2\log\left(\frac{1}{p}\right)} \qquad (1\text{-}29)$$

The same argument can be refined as follows: if $\zeta : p \mapsto 1 - \frac{q(p)}{\sqrt{2\log\left(\frac{1}{p}\right)}}$, the same process leads to $\zeta(p) = o\left(\sqrt{\log\left(\frac{1}{p}\right)}\right)$, and then that $q = \sqrt{2\log\left(\frac{1}{p}\right)} + o\left(\sqrt{\log\left(\frac{1}{p}\right)}\right)$.

With $m = q_{\min} \sim \sqrt{2\log\left(\frac{1}{p_M}\right)}$, $M = q_{\max} \sim \sqrt{2\log\left(\frac{1}{p_m}\right)}$, and $p = \varepsilon r$, the equation $\delta = 0$ is equivalent to $\sqrt{2\log\left(\frac{1}{\varepsilon r_M}\right)} = \rho\sqrt{2\log\left(\frac{1}{\varepsilon r_m}\right)}$. Let us denote by ε^* the ε that is the solution of this equation, it comes $\left(\frac{1}{\varepsilon^* r_m}\right)^{\rho^2} = \frac{1}{\varepsilon^* r_M}$ and then:

$$\varepsilon^* = \left(\frac{{r_m}^{\rho^2}}{r_M}\right)^{\frac{1}{1-\rho^2}} \tag{1-30}$$

Since $\rho \in [-1,1]$ and ρ can be interpreted as an angle in a representation where the risks are vectors, it is natural to define:

$$\theta \equiv \arccos(\rho) \in [0, \pi] \tag{1-31}$$

It is useful to notice that ρ intervenes only through ρ^2: positive and negative values of ρ (correlation vs anticorrelation) play a symmetric role in ε^*. In the empirical implementation used in this article, the exposure-based construction of Σ_E mostly produces non-negative correlations, because it is built from shared country/host exposure. Anticorrelations are therefore discussed here for mathematical completeness and for possible extensions, not because the present dataset is expected to generate many negative pairwise dependencies.

Moreover, the scale of the risk indicator is arbitrary, only the difference of ratio matters. Therefore, it is natural to capture this ratio, by defining:

$$\kappa \equiv \log\left(\frac{r_M}{r_m}\right) \in \mathbb{R}^+ \tag{1-32}$$

Finally, it is natural to rather consider a probability as a threshold rather than a parameter ε^* which scale is affected by the arbitrary scale of the risk indicator. It is therefore sound to introduce:

$$p^* \equiv \varepsilon^* r_M \tag{1-33}$$

With this definition, p^* will signal the smallest probability order of magnitude that triggers the change of mode. With these definitions comes the following result:

$$\begin{aligned} p^* &= r_M \exp\left(\frac{1}{\sin^2(\theta)}\left[\cos^2(\theta)\log(r_m) - \log(r_M)\right]\right) \\ &= \exp\left(\tan^{-2}(\theta)\log(r_m) + \left(1 - \frac{1}{\sin^2(\theta)}\right)\log(r_M)\right) \\ &= \exp(\tan^{-2}(\theta)\log(r_m) - \tan^{-2}(\theta)\log(r_M)) \end{aligned}$$

$$p^* = e^{-\kappa \tan^{-2}(\theta)} \tag{1-34}$$

A heatmap featuring p^* with the respect of θ and κ is displayed on Figure 3, with contour lines indicating the values of the probability threshold.

If $p \gg p^*$, $\overline{\Phi}(p, \theta, \kappa) \approx 1$ and then appears the "max mode" that was intuited:

$$R_{i \oplus j} \approx R_i + R_j - R_m = R_M \tag{1-35}$$

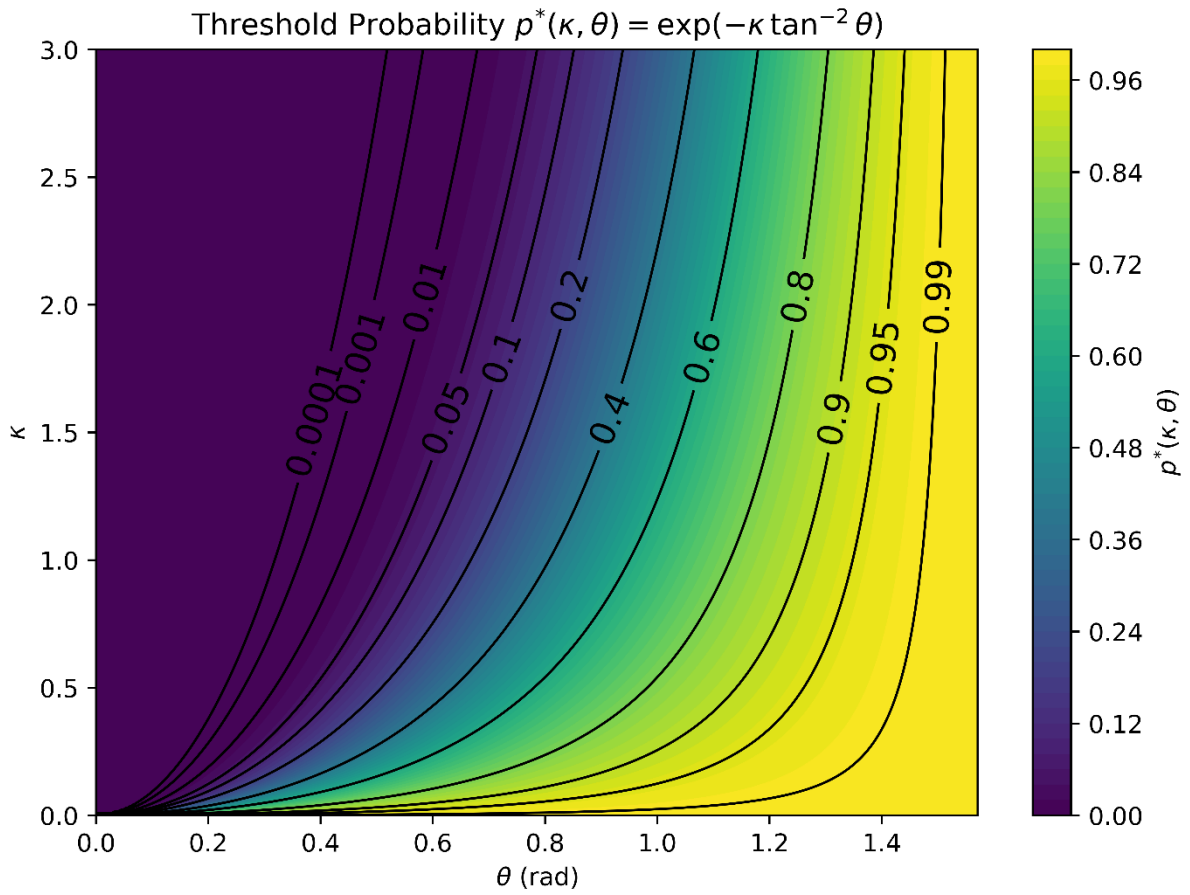

Figure 3: Threshold Probability p^* versus the Correlation between Elements i and j ($\cos(\theta)$) and the Log Ratio between the Intensity of their associated Risk (κ)

The threshold probability p^* associated with risks i and j gives the order of magnitude above which correlations between the two elements can no longer be neglected. When the incident probability p is below p^*, the correction factor $\bar{\Phi}(p,\theta,\kappa)$ – where $\theta \equiv \arccos(\rho_{ij})$ and $\kappa \equiv log(R_M/R_m)$ – tends to zero and the aggregation remains in the sum-like regime, with the union probability proportional to $(R_i + R_j)$. When p exceeds p^*, the correction factor becomes non-negligible and the aggregation moves toward a dependence-dominated regime.

Anticorrelations ($\theta > \frac{\pi}{2}$) play a role absolutely symmetric to correlations on p^*; in particular, the condition $p \gg p^*$ is likely to be fulfilled if elements E_i and E_j are fully anticorrelated.

What about δ, the parameter in function $\bar{\Phi}$? It depends of ε, and by defining $p = \varepsilon r_M$, by noticing that $\theta \in [0,\pi] \Rightarrow |\sin(\theta)| = \sin(\theta)$,

$$\delta(\varepsilon) = \frac{m - M\rho}{\sqrt{1-\rho^2}} \approx \frac{1}{\sin(\theta)}\left[\sqrt{2\log\left(\frac{1}{p}\right)} - \cos(\theta)\sqrt{2\log\left(\frac{r_M}{pr_m}\right)}\right]$$

$$= \frac{\sqrt{2\log\left(\frac{1}{p}\right)}}{\sin(\theta)}\left[1 - \cos(\theta)\sqrt{1 + \frac{\kappa}{\log\left(\frac{1}{p}\right)}}\right]$$

$$\delta(p) = \frac{\sqrt{2}}{\sin(\theta)}\left[\sqrt{\log\left(\frac{p^*}{p}\right) + \kappa\tan^{-2}(\theta)} - \cos(\theta)\sqrt{\kappa + \log\left(\frac{p^*}{p}\right) + \kappa\tan^{-2}(\theta)}\right] \quad (1\text{-}36)$$

One remaining question is how rapidly the aggregation moves from the sum-like regime to the dependence-dominated regime when p crosses the threshold $p \approx p^*$? To answer this question, let's notice that ρ must be big enough for the correlations to activate the transition between both modes, so $\theta \approx 0$ and $\sqrt{\tan^{-2}(\theta)} = \tan^{-1}(\theta)$, and let $\tau \equiv \log\left(\frac{p}{p^*}\right) \ll 1$.

$$\delta(\tau) = \frac{\sqrt{2\kappa}}{\sin(\theta)}\left[\tan^{-1}(\theta)\sqrt{1 - \frac{\tau}{\kappa\tan^{-2}(\theta)}} - \tan^{-1}(\theta)\sqrt{1 - \frac{\tau\sin^2(\theta)}{\kappa}}\right]$$

$$= \frac{\sqrt{2\kappa}\tan^{-1}(\theta)}{\sin(\theta)}\left[1 - \frac{\tau}{2\kappa\tan^{-2}(\theta)} - 1 + \frac{\tau\sin^2(\theta)}{2\kappa} + o(\tau)\right]$$

$$= \frac{-\tan^{-1}(\theta)}{\sqrt{2\kappa}\sin(\theta)}\frac{\sin^4(\theta)}{\cos^2(\theta)}\tau + o(\tau)$$

$$\boxed{\delta(\tau) = \frac{-\sin^2(\theta)}{\sqrt{2\kappa}\cos(\theta)}\tau + o(\tau)} \quad (1\text{-}37)$$

The values taken by $\bar{\Phi}(\delta)$ versus p and θ are displayed on heatmaps for two different values of κ on **Figure 4**. Significant corrections to the "mere sum" of indicators are necessary in the green-yellow part of the heatmaps.

Several factors prevent $\bar{\Phi}(\delta)$ from increasing quickly from $\bar{\Phi}(\delta = 0) = 1/2$ to $\bar{\Phi}(\delta) \approx 1$ ("max mode"):

- Reaching the threshold $p \geq p^*$ means having large κ and a small θ (the area where the values of p^* are the smallest on Figure 3). These conditions are precisely limiting the variations of δ: $\delta = O\left(\frac{\sin^2(\theta)}{\sqrt{\kappa}}\right)$.
- Since $\tau = \log\left(\frac{p}{p^*}\right)$, p must be significantly larger than p^* to make the first order in τ non-negligible.
- Last but not least, δ plays a role through $\bar{\Phi}(\delta)$, whereas $\bar{\Phi}'(0) = 0$.

As a conclusion, there is no sudden switch from the "sum mode" to the "max mode"; on the contrary, $\bar{\Phi}(\delta)$ varies slowly when p passes the threshold probability p^*. This indicates that

even p cannot be precisely assessed, being able to compare its probable order of magnitude with p^* already gives a reliable indication on the "mode". This also shows that there is no brutal switch from the "sum mode" to the "max mode", there exists a "transition mode" corresponding to p having a comparable order of magnitude to p^*, where the global risk indicator $R_{i\oplus j}$ is:

$$|\tau| \leq 1 \Longrightarrow R_{i\oplus j} \approx r_M + \frac{1}{2} r_m \quad (1\text{-}38)$$

In case of anticorrelation between E_i and E_j, $\cos(\theta) < 0$ and both terms in δ, $\sqrt{2\log\left(\frac{1}{p}\right)}$ and $\sqrt{2\log\left(\frac{r_M}{p r_m}\right)}$, add up, leading $\overline{\Phi}(\delta)$ to tend to 0, corresponding to a sum mode.

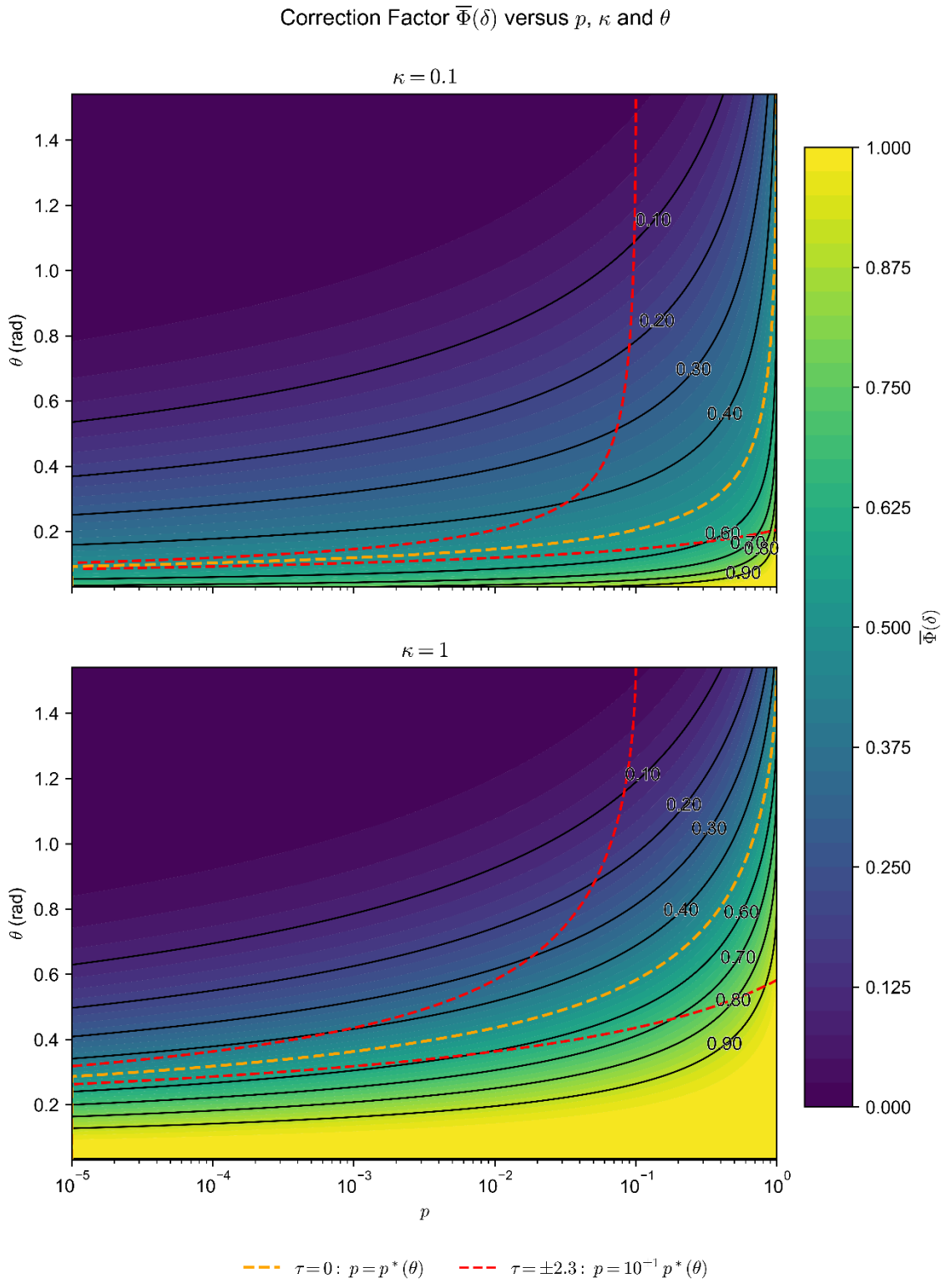

Figure 4: Correction factor to additive aggregation as a function of element correlation and risk-ratio contrast

This graph is based on the approximation of small probabilities p; values corresponding to $p >$ 0.1 are likely irrelevant.

The dash line in orange materializes the threshold probability $p^* = e^{-\kappa \tan^{-2}(\theta)}$ that is such that $\bar{\Phi}(p = p^*) = 1/2$. Above this line, $p < p^*$, and below, $p > p^*$. The dash lines in orange mark the area where p and p^* are of the same order of magnitude: $|\log_{10}(p/p^*)| \leq 1$. As explained above, the correction factor $\bar{\Phi}(\delta)$ keeps values close to $1/2$ inside this band.

Most of the graph is covered with small values of $\bar{\Phi}(\delta)$, corresponding to a sum-like regime, except when θ is small (strong correlations between the elements E_i and E_j) and the incident probability p is greater than the threshold probability p^*. As a conclusion, in absence of information of high correlations or high probabilities of incidents, defining $R_{\text{alloy}} \equiv R_i + R_j$ is a decent approximation.

Finally, it is interesting to consider the proportion of the error induced by the approximation ignoring the correction factor $R_{alloy} \approx r_{min} + r_{max}$, that is to say the ratio:

$$\left| \frac{\bar{\Phi}(\delta) r_{min}}{r_{min} + r_{max}} \right| = \left| \frac{\bar{\Phi}(\delta) e^{-\kappa}}{1 + e^{-\kappa}} \right| \leq e^{-\kappa} \tag{1-39}$$

Figure 3 shows that when θ is small and κ is large, the condition $p \gg p^*$ is likely. However, the relative correction to the rough "sum approximation" is bounded by a decreasing function in κ: $\left| \frac{R_i + R_j - R_{i \oplus j}}{R_i + R_j} \right| \leq e^{-\kappa}$. This ensures that, except for highly degenerated case for which $\theta \approx 0$ and then $p \gg p^*$ independently of the value of κ, the error remains moderate.

In conclusion, the parameter ε (or equivalently p) cannot be eliminated, but its order of magnitude can be compared with p^* to determine whether the aggregation remains close to the additive regime or enters a dependence-dominated regime. When no reliable estimate is available, a rough choice of p still allows one to define a combined indicator $R_{i \oplus j}$. The derived formulas formalize the initial intuition by providing an explicit threshold and a smooth transition between both regimes. Once the correlation matrix is known and a value of p is fixed, the extension of the elemental risk indicator $\boldsymbol{R}$ to alloys becomes straightforward and can be implemented numerically. If necessary, a simpler easy-to-interpret binary (rougher) approximation can even be defined by interpreting p_{ij}^* as a distance (defining $\theta_{ij} > \pi/2 \Longrightarrow d_{ij} \equiv 1$) between elements E_i and E_j in a clustering algorithm (e.g. linkage dendrogram cut with a threshold parameter p representing the assessed probability of incident):

$$R_{\text{alloy}} \approx \sum_{\mathcal{C} \text{ in clusters}} \max_{E_i \in \mathcal{C}} R(E_i) \tag{1-40}$$

To conclude this subsection, additive aggregation should be regarded as the default baseline in the absence of evidence for strong dependence or for sufficiently large incident probabilities. The present framework identifies and quantifies the subset of cases in which this baseline ceases to be adequate. Because the same first-order mapping is applied to all elements, pure elements and multi-element systems can be placed on a common comparative scale, enabling qualitative ranking across candidates with different numbers of constituents. The following formulas have been established:

$$R_{\text{alloy}} = \bigoplus_{E_i \in \text{ alloy}} R_i \equiv \sum_{E_i \in \text{ alloy}} R(E_i) - \sum_{\substack{i<j \\ \{E_i, E_j\} \in \text{ alloy}}} \bar{\Phi}(\delta_{ij}) \cdot R_m(i,j) \tag{1-41}$$

$$\Phi: x \mapsto \frac{1}{\sqrt{2\pi}} \int_{t=-\infty}^{x} e^{-\frac{t^2}{2}} dt \,, \bar{\Phi}: x \mapsto 1 - \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{t=x}^{+\infty} e^{-\frac{t^2}{2}} dt$$

$$\delta_{ij} = \frac{\sqrt{2}}{\sin(\theta_{ij})} \left[\sqrt{\log\left(\frac{1}{p}\right)} - \cos(\theta_{ij}) \sqrt{\kappa_{ij} + \log\left(\frac{1}{p}\right)} \right] \tag{1-42}$$

$R_M(i,j) \equiv \max(R_i, R_j)$, $R_m(i,j) \equiv \min(R_i, R_j)$

$\kappa_{ij} \equiv \log\left(\frac{R_M(i,j)}{R_m(i,j)}\right)$

$\theta_{ij} \equiv \arccos(\rho_{ij})$, with $\rho_{ij} \equiv (\Sigma_E)_{ij}$

$p_{ij}^* \equiv e^{-\kappa_{ij} \tan^{-2}(\theta_{ij})}$

$$\bar{\Phi}(\delta_{ij}) \approx \begin{cases} 0 \text{ if } p < p_{ij}^* \\ 1/2 \text{ if } p \approx p_{ij}^* \\ 1 \text{ if } p > p_{ij}^* \end{cases}$$

$$R_{i \oplus j} \approx \begin{cases} R_i + R_j \text{ if } p < p_{ij}^* \\ R_M(i,j) + R_m(i,j)/2 \text{ if } p \approx p_{ij}^* \\ R_M(i,j) \text{ if } p > p_{ij}^* \end{cases} \tag{1-43}$$

$$R_{\text{alloy}} \approx \sum_{\mathcal{C} \text{ in clusters}} \max_{E_i \in \mathcal{C}} R(E_i),$$

with a clustering method based on the distance $d(E_i, E_j) \equiv \begin{cases} p_{ij}^* \text{ if } \theta_{ij} \leq \frac{\pi}{2} \\ 1 \text{ if } \theta_{ij} > \frac{\pi}{2} \end{cases}$ (1-44)

This section formalizes how an elemental risk score can be extended to alloys and other multi-input systems while preserving a common comparative scale. The same logic can in principle be applied to more complex devices, such as smartphones, batteries, or broader material systems, whenever the objective is to compare the global supply-risk exposure of competing technologies or chemistries on the basis of their required elemental inputs. Applying the framework to full products would additionally require an explicit system boundary and, where relevant, quantity weights, substitutability assumptions, and recycling or secondary-supply pathways.

2. Technical construction of the dependence matrix from co-production and country exposure.

Section 1 shows that alloy aggregation requires an element–element dependence structure, encoded by the correlation coefficients ρ_{ij} (and thus $\theta_{ij} = \arccos(\rho_{ij})$ in Eq. (1-42)). The

purpose of Section 2 is to construct $\Sigma_E = (\rho_{ij})$ from objective supply-chain couplings (companionality and country-of-production).

Historically, metallurgy focused on a small set of 'major metals' – iron, copper, tin and lead – present at relatively high crustal abundances and extractable with comparatively simple processes (Nassar et al., 2015). Modern technologies, however, rely increasingly on 'minor metals' and critical elements (rare earths, germanium, indium, rhenium, etc.) used in semiconductors, superalloys, turbines, display technologies and medical imaging. Many of these elements occur at trace levels (often below 0.01 wt.%) and are not mined as primary products; they are recovered as by-products of mining operations targeting a host metal.

The concept of companionality, introduced by Nassar et al. (Nassar et al., 2015), quantifies this dependence: it is the degree to which a metal is produced as a by-product of one or more host metals. More than one-third of the metals used in modern technologies are produced at least 50% as by-products, often from geographically concentrated mining sectors.

A frequently cited visual representation is the "Wheel of Metals Companionality", which places host metals (e.g., nickel, copper, lead, molybdenum) in the inner ring and their dependent by-products (e.g., tellurium, indium, rhenium) in the outer ring, proportionally sized by their degree of dependence.

The historical trajectory of certain metals illustrates this phenomenon. Rhenium (Re), for instance, is produced almost exclusively as a by-product of molybdenum smelting, itself largely derived from copper porphyry deposits (Keller and Anderson, 2018). With an average crustal abundance of ~1 ppb, exceptional physical properties (very high melting point, mechanical strength, and chemical stability), and strategic uses in turbine blade superalloys and catalysts, rhenium has become indispensable in aerospace and other advanced industries.

However, its production is tightly coupled to the extraction of its host metals, leading to price inelasticity of supply: even if buyers were willing to pay twice the current price of rhenium, this would not double its production unless molybdenum and copper extraction rates also doubled. Unlike primary metals whose production can respond to price incentives, by-product metals exhibit low supply elasticity.

Companionality induces structural correlations in supply risks: any disruption to the production of a host metal translates directly into availability constraints for its by-products. For example, a copper mine strike or a catastrophic event reducing output will simultaneously impact the supply of tellurium, indium, and rhenium. Similarly, in the platinum group metals (PGM), rhodium, iridium, and ruthenium are almost entirely obtained as by-products of platinum mining in South Africa's Bushveld Complex, resulting in a nearly deterministic co-movement of their production volumes.

In the context of companionality, metals can be classified according to their production pathway. A *host metal* is extracted primarily for its own economic value, whereas a *by-product* is recovered incidentally during the extraction and processing of other metals. **Figure 5** provides a schematic example, using fictitious values, of the production dependencies among six elements: cobalt (Co), copper (Cu), iron (Fe), nickel (Ni), molybdenum (Mo), and rhenium (Re).

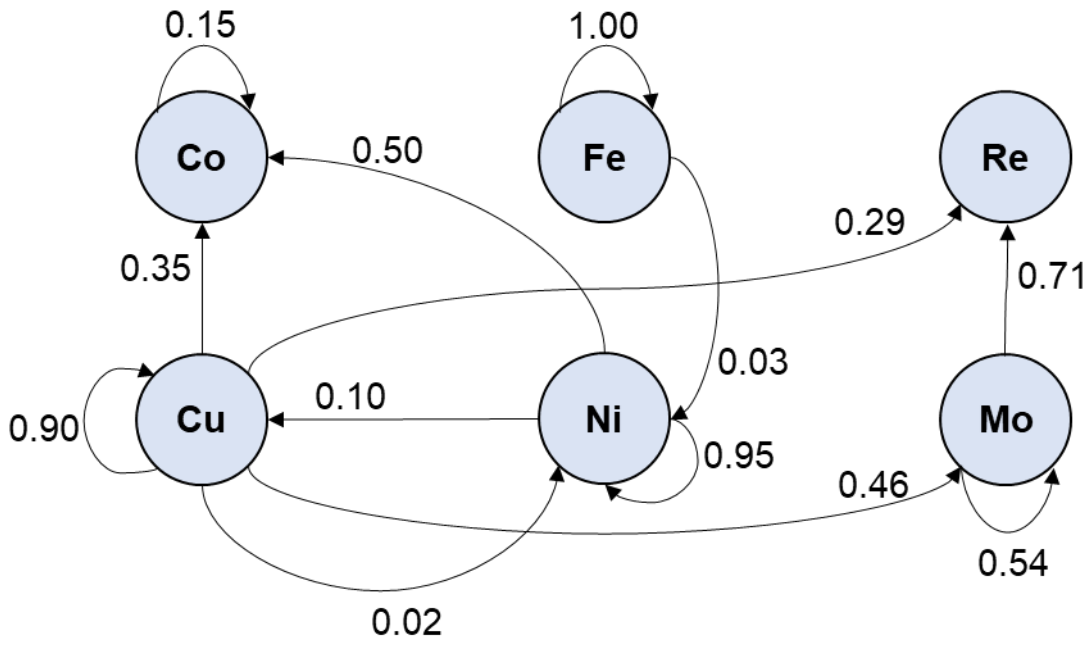

Figure 5: Fictitious companionality graph used to illustrate host–by-product propagation.

Each node represents an element. An arrow from node i to node j indicates that element j is partially obtained as a by-product of element i. The label on each arrow gives the fraction of j's production sourced from i. Loops on a node indicate the fraction of production obtained as a *primary product* (direct extraction for its own sake). For example, the loop on Cu (0.90) means that 90% of copper production is primary, while the arrow Cu → Co (0.35) means that 35% of cobalt is obtained from copper production.

Let us denote by $C \in \mathbb{R}^{n\times n}$ the *companionality matrix*, in which each column j describes the sourcing structure of element E_j from all potential "host" elements E_i. Specifically, the entry C_{ij} represents the fraction of the total production of element E_j that originates from the mining and processing of element E_i. By construction:

$$\forall i,j, 0 \leq C_{ij} \leq 1 \text{ and } \sum_{i=1}^{n} C_{ij} = 1 \qquad (2\text{-}1)$$

A nonzero diagonal entry C_{jj} indicates the proportion of primary production for E_j (direct extraction for its own sake), whereas the off-diagonal entries C_{ij} capture the by-product relationships.

In the present example, C is given by:

$$C = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.15 & 0 & 0 & 0 & 0 \\ 0.71 & 0 & 0.54 & 0 & 0 & 0 \\ 0 & 0.50 & 0 & 0.95 & 0.10 & 0 \\ 0.29 & 0.35 & 0.46 & 0.02 & 0.90 & 0 \\ 0 & 0 & 0 & 0.03 & 0 & 1 \end{pmatrix}$$

where the ordering of elements has been chosen such that the diagonal terms are in increasing order: Re, Co, Mo, Ni, Cu and Fe (n = 6). This ordering will be useful in subsequent developments.

2.1. From companionality graphs to Markov chains

Cascading and crossed by-product relationships can be represented compactly using an absorbing Markov-chain formulation.

A Markov process is a discrete-time stochastic process $\{M_t\}_{t\geq 0}$ taking values in a finite or countable state space S, for which predictions can be made regarding future outcomes based solely on its present state. Its future depends upon the present only, it is not influenced by its past:

$$\mathbb{P}(M_{t+1} = s_{t+1} | M_t = s_t \dots M_0 = s_0) = \mathbb{P}(M_{t+1} = s_{t+1} | M_t = s_t)$$

The transition dynamics are described by a **transition matrix** $T = (T_{ij})$ where:

$$T_{ij} = \mathbb{P}(M_{t+1} = s_j | M_t = s_i), \sum_j T_{ij} = 1, 0 \leq T_{ij} \leq 1 \qquad (2\text{-}2)$$

A state $s \in S$ is called **absorbing** if:

$$\mathbb{P}(M_{t+1} = s | M_t = s) = 1 \qquad (2\text{-}3)$$

In other words, once the process enters s, it remains there forever. A Markov chain is said to be absorbing if:

- There exists at least one absorbing state;
- From every state in the chain, there is a nonzero probability of eventually reaching an absorbing state in a finite number of steps.
- The states that are not absorbing are called “transient”.

In this setting, states correspond to chemical elements; the companionality matrix C can be interpreted as a *column-stochastic* transition matrix; absorbing states correspond to elements that are 100% “primary” ($C_{ii} = 1$): once the production chain reaches such a state, no further “upstream” sourcing occurs; **t**ransient states correspond to elements that are at least partially produced as by-products ($C_{ii} < 1$): these states eventually lead to one or more absorbing states via successive by-product relationships.

Once it is noticed that element E_{i_1} can be a by-product of element E_{i_2}, E_{i_2} being himself a by-product of E_{i_3}, the fact that there are domino-effect in being a by-product is obvious: E_{i_1} being a direct by-product of E_{i_2} implies it is a "2-steps" by-product of E_{i_3} and so on. The k-steps domino effect is given by C^k, and if element E_i is considered, its host elements in exactly k steps are given by the values of $C^k \boldsymbol{e_i}$ where $\boldsymbol{e_i}$ is the i-th canonical base vector.

The number of transient states (t_i) will be denoted by u, and the number of absorbing states (a_i) by v; $u + v = 5 + 1 = n$, and the elements are reordered to ensure that the last v are absorbing. Then the companionality matrix can be written:

$$C = \begin{pmatrix} T & 0_{u\times v} \\ S & I_v \end{pmatrix} \tag{2-4}$$

Where T is a $u \times u$ matrix which holds probabilities of moving from a by-product to a by-product state, and S is a $v \times u$ matrix which holds the probabilities of movigin from a by-product to a 100% primary state. Physically, any by-product must come from a host, so the eventual goal is to know the matrix C^k when k is big enough, which means $\lim_{k\to+\infty} C^k$: one can now show that this limit exists and how to compute it.

By using the formula for the product of matrices by blocs, it comes:

$$C^k = \begin{pmatrix} T^k & 0_{u\times v} \\ S\sum_{l=0}^{k-1} T^l & I_v \end{pmatrix} \tag{2-5}$$

Does the series $\sum_{l=0}^{k-1} T^l$ converge? Physically, it must, because any by-product must indirectly come from primary elements. This means that S cannot be 0 – and that there is a "loss" from transient states to absorbing states at every step. More precisely, there is no "recursive closed loop": there cannot be a subset of t in which every element is a by-product of elements of this subset of t. This implies that:

$$\forall t_i \in t, \exists k_i, k \geq k_i \Rightarrow \sum_{l=1}^{u} (T^k)_{li} < 1 \tag{2-6}$$

and by defining $K \equiv \max_{i\in[\![1,u]\!]} k_i$,

$$\forall i \in [\![1,u]\!], k \geq K \Rightarrow \sum_{l=1}^{u} (T^k)_{li} \leq \xi \equiv \max_{i\in[\![1,u]\!]} \sum_{l=1}^{u} (T^k)_{li} < 1 \tag{2-7}$$

From Perron-Frobenius theorem, it is known that the spectral radius $\rho(T^K)$ of T^K is less than the maximum value of the sum of the values of every column, so $\rho(T^K) \leq \xi < 1$. From this comes:

$$\sum_{l=1}^{u} (T^k)_{li} \leq \xi < 1 \Rightarrow \rho(T^k) \leq \|T^K\|_1 \leq \xi < 1 \Longrightarrow \rho(T) = \rho(T^k)^{1/K} \leq \xi^{1/K} < 1 \quad (2\text{-}8)$$

This proves that the series $\{T^l\}_l$ converges, that $(I_u - T)$ is inversible, and that

$$\lim_{k \to +\infty} S \sum_{l=0}^{k-1} T^l = S \cdot (I_u - T)^{-1} \quad (2\text{-}9)$$

Since the series $\{T^l\}_l$ converges, it comes:

$$\lim_{k \to +\infty} T^k = 0_{u \times u} \quad (2\text{-}10)$$

And from all this, the following important result comes:

$$C^\infty \equiv \lim_{k \to +\infty} C^k = \begin{pmatrix} 0_{u \times u} & 0_{u \times v} \\ S \cdot (I_u - T)^{-1} & I_v \end{pmatrix} \quad (2\text{-}11)$$

This first construction is not yet sufficient. Indeed, the absorbing Markov chain model that was used implies that any production of E_i at step k will be split between the host products, whereas this is not the case: if 90% of Cu production comes from Cu-mines, these 90% should not be split, at the next step, between host elements. And this complexity propagates: if 46% of Mo comes from Cu, it means that 90% of 46% are related to Cu exploitation and should not be split further by the model.

2.2. Virtual absorbing states and construction of the element-to-host matrix $\aleph_{E/HE}$

To represent primary production correctly, we reconfigure the companionality graph by adding virtual absorbing states. This makes it possible to use an absorbing Markov-chain formulation while preserving primary-production shares. For this, one can add virtual absorbing summits (marked with a star) to describe primary production, as shown on **Figure 6**:

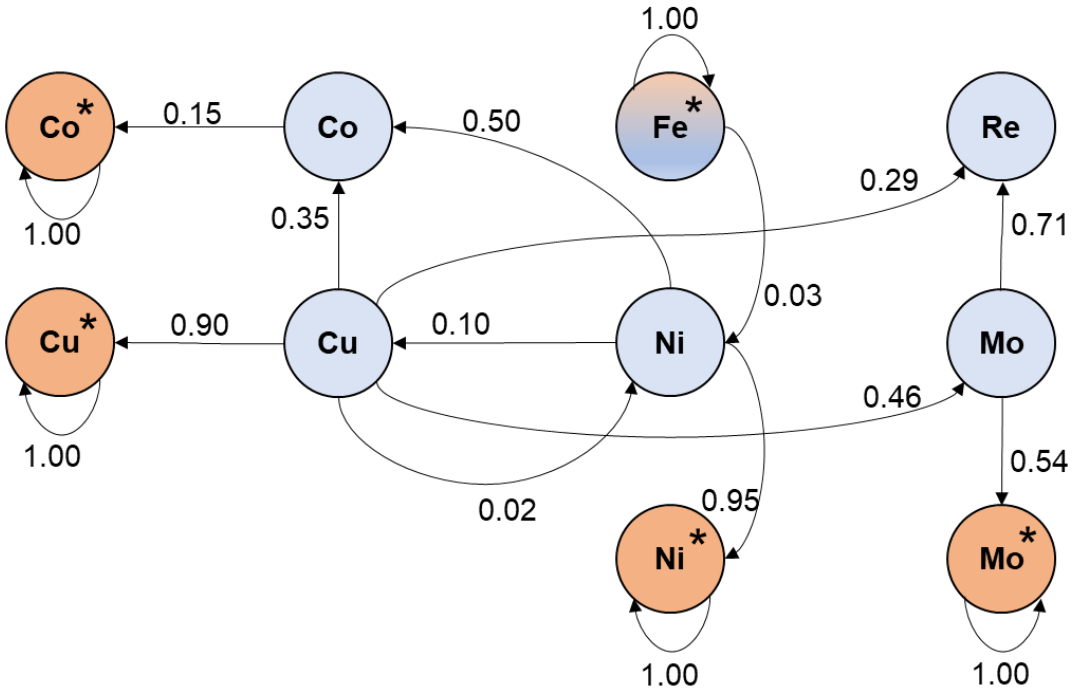

Figure 6: Companionality Graph Segregating Primary Production

To be able to use the absorbing Markov chains, the companionality graph is redesigned to make the primary production absorbing. Virtual absorbing states Cu^*, Co^*, Ni^*, Mo^* are added to make to model fit the fact that primary production is not to be split in host production. Fe being already absorbing, it does not need a new virtual state.

The transition matrix associated to this graph is of higher dimension: $u + u^* + v = 5 + 4 + 1 = 10$ instead of $u + v = 6$, u^* being the number of virtual states added (the number of non-zero diagonal term of T).

$$C^* = \begin{pmatrix} T^* & 0_{u\times u^*} & 0_{u\times v} \\ \mathrm{D} & I_{u^*} & 0_{u^*\times v} \\ S & 0_{v\times u^*} & I_v \end{pmatrix} = \begin{pmatrix} T^* & 0_{u\times(u^*+v)} \\ \binom{D}{S} & I_{u^*+v} \end{pmatrix} \qquad (2\text{-}12)$$

With T^* being T deprived from its diagonal terms $T^* = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0.71 & 0 & 0 & 0 & 0 \\ 0 & 0.50 & 0 & 0 & 0.10 \\ 0.29 & 0.35 & 0.46 & 0.02 & 0 \end{pmatrix}$,

and $D \in \mathcal{M}_{u^*\times u}(\mathbb{R})$ being the matrix holding the diagonal terms of T:

$$D = \begin{pmatrix} 0 & 0.15 & 0 & 0 & 0 \\ 0 & 0 & 0.54 & 0 & 0 \\ 0 & 0 & 0 & 0.95 & 0 \\ 0 & 0 & 0 & 0 & 0.90 \end{pmatrix}$$

And $S \in \mathcal{M}_{v\times u}(\mathbb{R})$ carries the flux from transient states to the original absorbing states:

$$S = (0 \quad 0 \quad 0 \quad 0.03 \quad 0)$$

The product $\binom{D}{S} \cdot (I_u - T^*)^{-1} \in \mathcal{M}_{(u^*+v)\times u}(\mathbb{R})$ – replacing former $S \cdot (I_u - T)^{-1}$ from the first graph – is very interesting: it gives, for any transient element X_j, the fraction that derives from the primary extraction of X_i.

$$\binom{D}{S} \cdot (I_u - T^*)^{-1} = \begin{pmatrix} 0 & 0.15 & 0 & 0 & 0 \\ 0.38 & 0 & 0.54 & 0 & 0 \\ 0.06 & 0.51 & 0.04 & 0.95 & 0.10 \\ 0.56 & 0.32 & 0.41 & 0.02 & 0.90 \\ 0 & 0.02 & 0 & 0.03 & 0 \end{pmatrix}$$

Which reads, in terms of deriving from primary extraction:

- Re derives at 38% from Mo, 6% from Ni and 56% from Cu.
- Co derives at 15% from Co, 51% from Ni, 32% from Cu and 2% from Fe
- Mo derives at 54% from Mo, 4% from Ni and 41% from Cu.
- Ni derives at 95% from Ni, 2% from Cu and 3% from Fe
- Cu derives at 10% from Ni and 90% from Cu.

It is to be noted that the matrix $(I_u - T^*)^{-1}$ counts the average number of "visits" to a transient state j, given that one started from element i. For later use, the matrix that associates to each element its ratio of host elements (E_i is now defined to be a "host element" if and only if $C_{ii} > 0$) is defined to be $\aleph_{E/HE} \in \mathcal{M}_{(u^*+v)\times n}(\mathbb{R})$:

$$\aleph_{E/HE} = \begin{pmatrix} \binom{D}{S} \cdot (I_u - T^*)^{-1} & 0_{u^*\times v} \\ & I_v \end{pmatrix} \qquad (2\text{-}13)$$

From now on, using the companionality matrix, the share of element deriving from the primary extraction of elements can be deduced, and therefore it is possible to propagate through $\aleph_{E/HE}$ a primary extraction related risk from host elements to all others. This matrix is directional: a companion element may depend strongly on a host metal, without the reverse being true. The directionality is essential for propagating host-country exposure to companion elements. In contrast, the correlation matrix used later in the Gaussian copula is symmetric by construction and should be interpreted as a pairwise measure of shared exposure, not as a causal map of host–by-product dependence. An illustration of this on PGM elements is done on **Table 1**.

	Ru	**Rh**	**Re**	**Ir**	**Pd**	**Pt**
Pd					3,0%	
Ag			0,2%			
Mo			38,3%			
Pt	80,1%	71,9%	0,6%	80,1%	38,0%	84,3%
Au			0,1%			
Zn			0,7%			
Pb			0,7%			
Cu			56,3%			
Ni	19,9%	28,1%	3,1%	19,9%	59,0%	15,7%

Table 1: Extract of Matrix $\aleph_{E/HE}$ relating PGM Elements to their Host Elements

This table is an extract of the whole matrix $\aleph_{E/HE}$ limited on the PGM elements: Ru, Rh, Re, Ir, Pd, Pt – Os was previously excluded from the elements set. Companions are represented as columns, while their hosts stand in lines. It reveals the "effective companionality" considering the indirect contributions. The full numerical matrix is provided as a CSV file in the online repository (Zenodo, DOI: 10.5281/zenodo.17573902) and mirrored on GitHub (https://github.com/fr55160/materials_inverse_design_pipeline/blob/main/scripts/G-risks/2-Supply%20Risk/M_E_over_HE.csv)

The supply risk profile of many critical raw materials (CRMs) is strongly influenced by the countries where they are extracted. Numerous studies have shown that geopolitical, economic, environmental, and social conditions in producing countries can significantly affect the stability of mineral supply chains. For example, Lèbre et al. (Lèbre et al., 2020) quantified the overlap between mineral extraction and high-risk contexts, showing that 70% of cobalt and 84% of platinum resources are located in jurisdictions with elevated environmental, social, and governance (ESG) risks. Similarly, Nassar et al. (Nassar et al., 2015) emphasized that country-specific supply disruptions – due to political instability, trade restrictions, or labor disputes – can propagate to multiple dependent metals via co-production relationships. In addition, the Fraser Institute Annual Survey of Mining Companies (Yunis, 2020) provides empirical evidence that policy-related factors, such as regulatory uncertainty, taxation, and infrastructure quality, are major determinants of investment attractiveness for the mining sector. Their Policy Perception Index (PPI) measures perceived policy risk, highlighting significant variation across jurisdictions.

Several indicators are consistently used in the literature to represent country-level risks relevant to CRM supply. **Table 2** summarizes the most notable ones, along with the type of risk conveyed and the sources that reference them.

INDICATOR	RISK NATURE	SOURCE(S)
POLICY PERCEPTION INDEX (PPI)	Regulatory and political risk	(Yunis, 2020)
RESOURCE GOVERNANCE INDEX (RGI)	Governance quality, transparency	NRGI, cited in (Lèbre et al., 2020)
WORLDWIDE GOVERNANCE INDICATORS (WGI)	Control of Corruption, Government Effectiveness, Political Stability and Absence of Violence/Terrorism, Rule of Law, Regulatory Quality, Voice and Accountability	World Bank, cited in (Cimprich et al., 2019; Gemechu et al., 2015; "Home | Worldwide Governance Indicators," 2025; Nassar et al., 2015)
ENVIRONMENTAL PERFORMANCE INDEX (EPI)	Environmental regulatory enforcement	(Lèbre et al., 2020)
HUMAN DEVELOPMENT INDEX (HDI)	Social development risk	(Cimprich et al., 2019; Gemechu et al., 2015; Koyamparambath et al., 2022; Lèbre et al., 2020)
CONFLICT MINERALS REPORTING	Armed conflict, human rights	OECD Guidance, cited in ESG literature
TRADE RESTRICTION INDEX	Tariff and export ban risk	(Nassar et al., 2015)
HERFINDAHL-HIRSCHAM INDEX (HHI)	Concentration / Monopoly of Production	(Cimprich et al., 2019; Nassar et al., 2015)
CORRUPTION PERCEPTIONS INDEX (CPI)	Governance and transparency	Transparency International, cited in (Lèbre et al., 2020)

Table 2: Main Risk Indicators Associated with CRM Supply Risk

Many risks usually considered in the criticality literature are country-related or country-mediated, even though supply risk also includes mechanisms that are not reducible to country exposure.

These proxies enable a quantitative approach to country risk evaluation. Some, such as the PPI or CPI, are perception-based, while others like the EPI or HDI are statistical composites.

The conclusion that "a significant portion of CRM supply risk derives from country-specific factors" is supported by multiple converging findings:

- Concentration in high-risk jurisdictions: Many CRMs are extracted predominantly in a few countries with high governance, political, or ESG risk. For example, cobalt production is heavily concentrated in the Democratic Republic of Congo, a country with low governance scores and high political instability.
- Co-production amplifies exposure: When by-products are co-extracted with primary metals in high-risk jurisdictions, the by-products inherit the primary metal's country risk profile.
- Lack of short-term substitution: For many CRMs, alternative supply sources or substitution technologies are limited, making country risk a persistent structural factor.

Once the country-level risks are quantified for primary metals, they can be propagated to by-products using the companionality by-products-to-hosts matrix $\aleph_{E/HE}$, and the matrix $\aleph_{HE/C}$ providing, for every host element, the shares of its production by country. Then the risk associated by country $\boldsymbol{R_C}$ propagates to elements in a vector $\boldsymbol{R_E}$:

$$\boldsymbol{R_E} = \aleph_{E/HE} \cdot \aleph_{HE/C} \cdot \boldsymbol{R_C} \qquad (2\text{-}14)$$

This combined matrix $\aleph_{E/HE} \cdot \aleph_{HE/C}$ can be interpreted as the effective production share of each country for each element, once companionality is considered. For diagnostic purposes, we distinguish two country-exposure matrices: the direct exposure matrix, built from element-country or host-country shares before companionality propagation, and the effective exposure matrix $\aleph_{E/HE} \cdot \aleph_{HE/C}$, obtained after propagating host–by-product dependencies. Comparing these two matrices isolates the role of companionality in reshaping geographical exposure before the final correlation matrix is computed. **Figure 7** illustrates these effective shares for the HEA candidate elements, highlighting the strong dominance of a small number of producing countries.

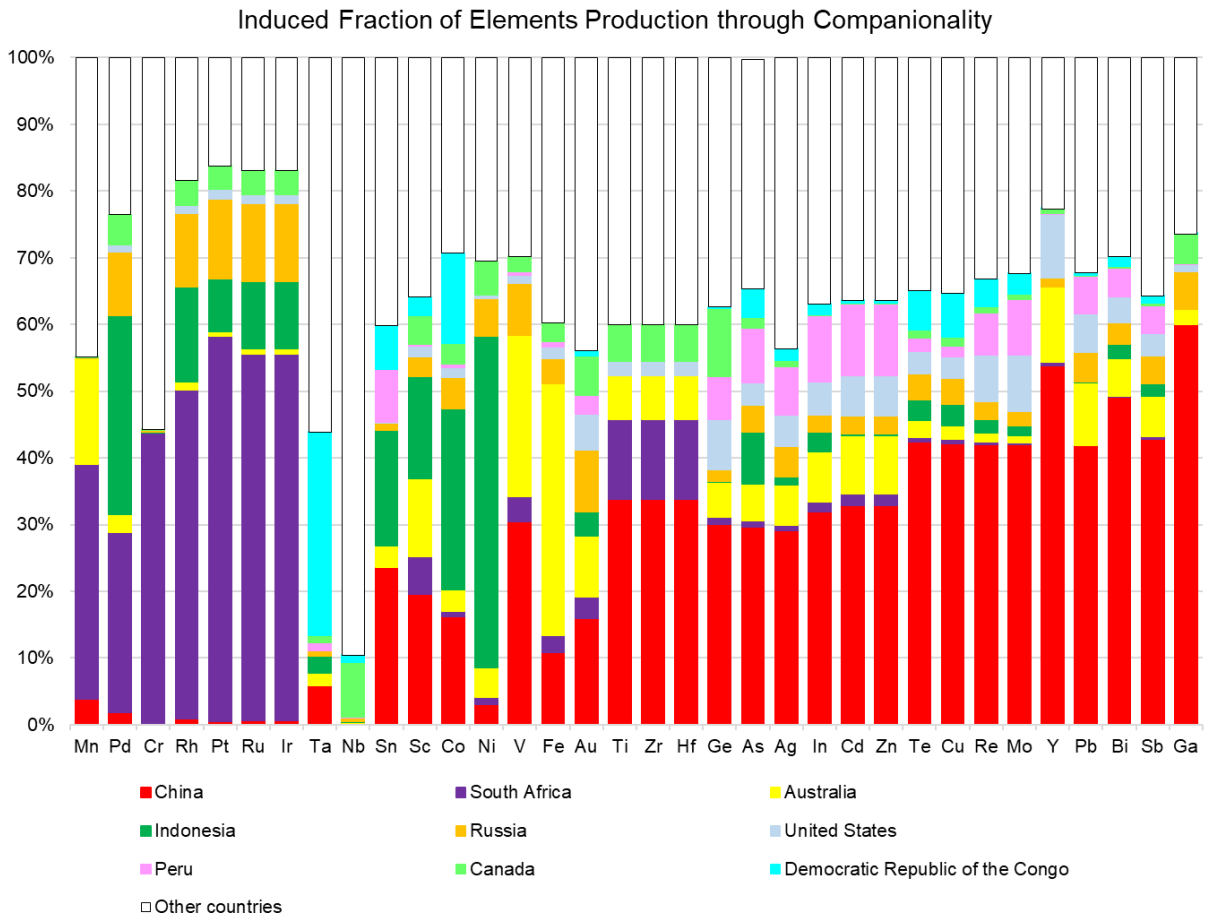

Figure 7: Effective country exposure of candidate elements after companionality propagation

The product $\aleph_{E/HE} \cdot \aleph_{HE/C}$ combines effective host contributions with host-country production shares and gives the effective country share in each element's supply. This matrix content is displayed as bars graph dealing with the metal candidates for HEA. For visual purposes, the data of 10 countries (out of 135 in the matrix) are explicated. The graphical representation shows that a limited number of countries account for most of the worldwide production, in particular China, and South-Africa for PGMs.

The shares of production per country were sourced from USGS (USGS, 2024).

To propagate correlations through country exposures within a Gaussian-copula framework, we introduce a latent Gaussian vector $\boldsymbol{Z_C} \sim \mathcal{N}(\mathbf{0}, I_C)$ with independent country components. This is not an assumption that the country-risk scores $\boldsymbol{R_C}$ are Gaussian; it is a convenient latent representation to generate an element-level correlation structure through linear mixing. As a consequence, the Gaussian vector associated to elements $\boldsymbol{Z_E} = \aleph_{E/HE} \cdot \aleph_{HE/C} \cdot \boldsymbol{Z_C} \sim \mathcal{N}\left(\mathbf{0}, \aleph_{E/HE} \cdot \aleph_{HE/C}\left(\aleph_{E/HE} \cdot \aleph_{HE/C}\right)^{\top}\right)$.

In this implementation, we model the dependence component of supply risk as mainly arising from shared country exposure after companionality propagation. Under this working assumption, the covariance matrix between elements $\mathrm{cov}(E)$ is

$\mathrm{cov}(E) = \aleph_{E/HE} \cdot \aleph_{HE/C}\left(\aleph_{E/HE} \cdot \aleph_{HE/C}\right)^{\mathsf{T}}$, and the correlation matrix Σ_E is:

$$\Sigma_E(i,j) = \frac{\mathrm{cov}(E)_{ij}}{\sqrt{\mathrm{cov}(E)_{ii}\mathrm{cov}(E)_{jj}}} \tag{2-15}$$

This Σ_E provides the coefficients $\rho_{ij} \equiv (\Sigma_E)_{ij}$ used in Eq. (1-42), hence in the threshold probabilities p_{ij}^* (Eq. (1-34)) and in the correction factor $\overline{\Phi}(\delta_{ij})$ entering the alloy aggregation formula (Eq. (1-41)).

The correlation matrix shown below is therefore the final combined result of two steps: country exposure and its redistribution through companionality. It should be interpreted together with the direct/effective exposure comparison above. The resulting element–element correlation structure is summarized by the correlation matrix Σ_E. To make these patterns readable, **Figure 8** shows the associated heatmap, with elements reordered according to a hierarchical clustering so that groups of strongly correlated elements appear as contiguous blocks.

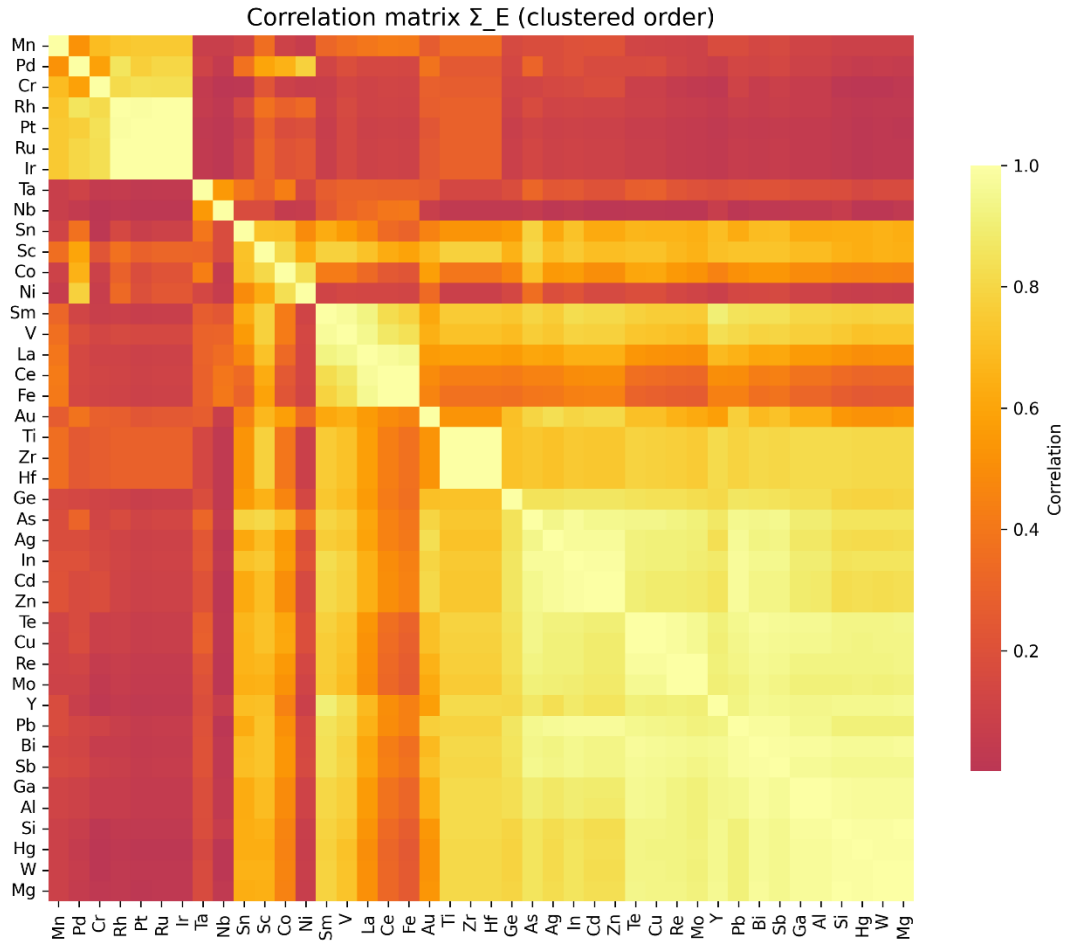

Figure 8: Element–element correlation matrix $\boldsymbol{\Sigma_E}$ induced by shared country exposure and companionality

The underlying correlation matrix, containing thousands of entries, is difficult to inspect directly and is therefore not printed here. It is provided in full, in machine-readable form, in the open data and code repository associated with this study (see Data availability section of the main article). The heatmap reveals clusters of elements with high shared exposure under the

present model (for readability, elements were reordered according to the leaves of a dendrogram associated to the distance $d(i,j) = 1 - \mathrm{corr}(i,j)$). Since this is a symmetric correlation matrix, these clusters should not be interpreted as directional host–by-product relationships; the latter are encoded in the companionality and effective host-contribution matrices.

These clusters indicate that additive aggregation may overestimate system-level risk for subsets of strongly dependent elements. For many other element pairs, however, the threshold analysis remains consistent with additive aggregation as a useful first approximation. This is in particular due to the fact that few countries produce most of the quantities (PGMs in South-Africa, most others in China; see **Figure 7** were the order of elements is the same). The risk is therefore strongly correlated to the risk conveyed by these countries.

The shares of production per country were sourced from USGS (USGS, 2024).

2.3. Application to the European Commission 2023 Supply Risk indicators

The European Commission's 2023 assessment of critical raw materials (Grohol and Veeh, n.d.) provides the Supply Risk (SR) indicator used as the worked example in the main article. The SR indicator is a composite, dimensionless score designed for cross-material comparison. It combines supplier concentration, EU sourcing where relevant, import reliance, governance and trade-related factors, recycling, substitution, and the value-chain stage at which the bottleneck occurs.

For the present article, SR is used only as the marginal elemental score. We do not attempt to recalibrate SR into an absolute disruption probability. Instead, the Supplementary Information derives a first-order score-to-probability mapping that preserves relative comparisons and allows multi-element aggregation to obey the probability laws of the underlying latent incident events.

This distinction is important. The marginal SR score retains the broader composite content of the European Commission indicator, whereas the dependence matrix constructed in this Supplementary Information captures only the co-movement channels induced by shared country exposure and co-production / companionality. The objective is therefore not to redefine EU criticality, but to build a system-level aggregation layer suitable for comparing multi-element systems with different numbers of constituent elements.

Therefore, to use & combine the supply risk indicators provided by the European Commission's 2023 assessment of critical raw materials assessment (European Commission. Directorate General for Internal Market, Industry, Entrepreneurship and SMEs., 2020) for the alloys of interest in this study, the calculation of the threshold probability given by Eq. (1-34) is used with the correlation matrix derived from countries of production & companionality (2-15). The result is presented in **Figure 9**.

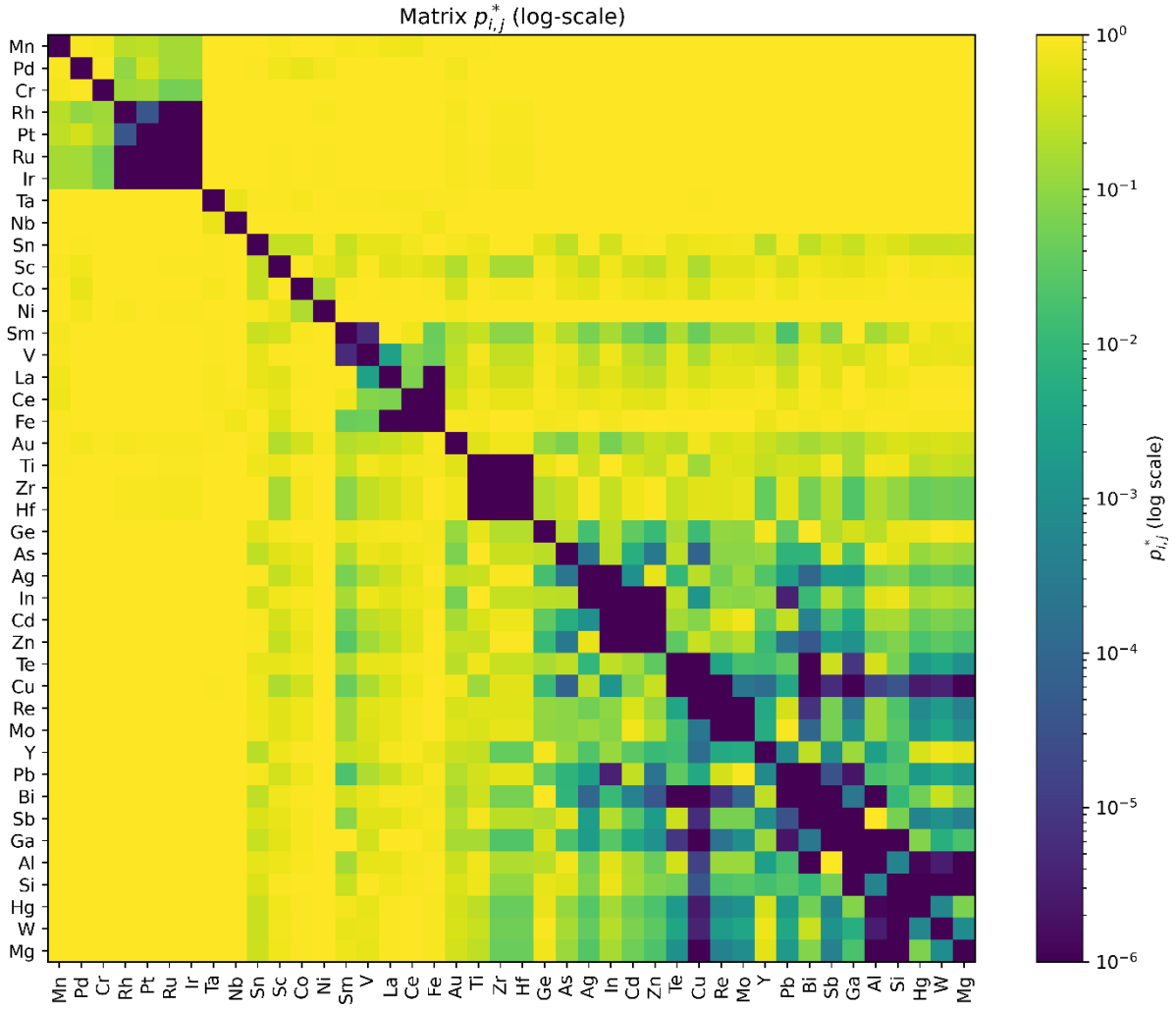

Figure 9 : Threshold Probabilities p^* as defined in Eq. (1-34) for the Supply Risk Indicators provided by the EU for CRM

Most threshold probabilities associated to a couple of elements are elevated (more than 0.1), meaning that the appropriate way to combine the supply risk indicators associated to these elements is a mere sum. However, some clusters of elements with significantly low threshold probability appear, requiring the use of formula (1-26) to combine rigorously the risk indicators.

A defensible working assumption for this quantitative framework is that the annual probability that a given critical raw material (CRM) experiences a "supply-risk incident" lies on the order of 10^{-3}-10^{-2}. This modest range is consistent with what can be inferred from external proxies and historical episodes, while acknowledging that most published "supply risk" metrics are relative scores rather than calibrated probabilities (Nassar et al., 2020). First, global monitoring shows that export restrictions affecting industrial raw materials have become widespread: about 14% of world trade flows in these materials faced at least one restriction over 2021-2023 (OECD, 2025). If interpreted as a Poisson arrival for "some restriction somewhere," that three-year incidence implies a generic annual hazard of $-\log(1-0.14)/3 \approx 5\%$ at the trade-basket level, which necessarily overstates the hazard for one specific metal and for incidents that materially impair EU access – hence the down-scaled 0.1-1%/year range adopted here. Second,

well-documented disruptions (e.g., the 2014 South African platinum strike halting ~40% of global Pt for five months (Stoddard, 2014); Indonesia's nickel ore export bans in 2014 and 2020 (Guberman et al., 2024); the 2010 rare-earths shock under Chinese controls) confirm that severe events do occur but are episodic at the commodity level (Sprecher et al., 2017) – supporting probabilities well below a few percent per metal-year, yet clearly above 10^{-4} given observed frequencies across many metals and years. Because the alloy-level aggregation is insensitive to small variations in p^*, using 10^{-3}-10^{-2} per year (thus $\varepsilon = 10^{-3}$ in Eq. (1-26)) as an order-of-magnitude prior is a pragmatic and conservative choice pending better calibration.

References

Cimprich, A., Bach, V., Helbig, C., Thorenz, A., Schrijvers, D., Sonnemann, G., Young, S.B., Sonderegger, T., Berger, M., 2019. Raw material criticality assessment as a complement to environmental life cycle assessment: Examining methods for product-level supply risk assessment. J of Industrial Ecology 23, 1226–1236. https://doi.org/10.1111/jiec.12865

European Commission. Directorate General for Internal Market, Industry, Entrepreneurship and SMEs., 2020. Study on the EU's list of critical raw materials (2020): final report. Publications Office, LU. https://doi.org/10.2873/11619

Gemechu, E.D., Helbig, C., Sonnemann, G., Thorenz, A., Tuma, A., 2015. Import-based Indicator for the Geopolitical Supply Risk of Raw Materials in Life Cycle Sustainability Assessments. J of Industrial Ecology 20, 154–165. https://doi.org/10.1111/jiec.12279

Grohol, M., Veeh, C., n.d. Study on the Critical Raw Materials for the EU 2023 - Final Report - European Commission [WWW Document]. https://doi.org/10.2873/725585

Guberman, D., Schreiber, S., Perry, A., 2024. Export Restrictions on Minerals and Metals: Indonesia's Export Ban of Nickel.

Home | Worldwide Governance Indicators [WWW Document], 2025. . World Bank. URL https://www.worldbank.org/en/publication/worldwide-governance-indicators (accessed 8.16.25).

Keller, P., Anderson, C., 2018. The Production of Critical Materials as By Products. Aspects in Mining & Mineral Science 2, 14. https://doi.org/10.31031/AMMS.2018.02.000532

Koyamparambath, A., Santillán-Saldivar, J., McLellan, B., Sonnemann, G., 2022. Supply risk evolution of raw materials for batteries and fossil fuels for selected OECD countries (2000–2018). Resources Policy 75, 102465. https://doi.org/10.1016/j.resourpol.2021.102465

Lèbre, É., Stringer, M., Svobodova, K., Owen, J.R., Kemp, D., Côte, C., Arratia-Solar, A., Valenta, R.K., 2020. The social and environmental complexities of extracting energy transition metals. Nature Communications 11, 4823. https://doi.org/10.1038/s41467-020-18661-9

Nassar, N.T., Brainard, J., Gulley, A., Manley, R., Matos, G., Lederer, G., Bird, L.R., Pineault, D., Alonso, E., Gambogi, J., Fortier, S.M., 2020. Evaluating the mineral commodity supply risk of the U.S. manufacturing sector. Sci. Adv. 6, eaay8647. https://doi.org/10.1126/sciadv.aay8647

Nassar, N.T., Graedel, T.E., Harper, E.M., 2015. By-product metals are technologically essential but have problematic supply. Sci. Adv. 1, e1400180. https://doi.org/10.1126/sciadv.1400180

OECD, 2025. OECD Inventory of Export Restrictions on Industrial Raw Materials 2025: Monitoring the use of export restrictions amid growing market and policy tensions. OECD Publishing. https://doi.org/10.1787/facc714b-en

Sprecher, B., Daigo, I., Spekkink, W., Vos, M., Kleijn, R., Murakami, S., Kramer, G.J., 2017. Novel Indicators for the Quantification of Resilience in Critical Material Supply Chains, with a 2010 Rare Earth Crisis Case Study. Environ. Sci. Technol. 51, 3860–3870. https://doi.org/10.1021/acs.est.6b05751

Stoddard, E., 2014. South Africa miners return to work after longest platinum strike. Reuters.

USGS, 2024. Mineral commodity summaries 2024 (Report No. 2024), Mineral Commodity Summaries. USGS, Reston, VA. https://doi.org/10.3133/mcs2024

Yunis, J., 2020. Fraser Institute Annual Survey of Mining Companies 2020.