Supplementary Materials for

Evidence for a Low Bulk Crustal Density for Mars from Gravity and Topography

Sander Goossens, Terence J. Sabaka, Antonio Genova, Erwan Mazarico,
Joseph B. Nicholas, Gregory A. Neumann

July 17, 2017

Contents

1 Rank minus one (RM1) constraint

We seek to minimize a cost function $J(\mathbf{x})$ with respect to model parameter vector $\mathbf{x}$ given by

$$J(\mathbf{x}) = \frac{1}{2}\left(\mathbf{d}-\mathbf{a}(\mathbf{x})\right)^{\mathrm{T}}\mathbf{W}\left(\mathbf{d}-\mathbf{a}(\mathbf{x})\right) + \frac{1}{2}\lambda\left(\mathbf{x}_a-\mathbf{x}\right)^{\mathrm{T}}\mathbf{P}\left(\mathbf{x}_a-\mathbf{x}\right), \tag{1}$$

where $\mathbf{d}$ is a data vector, $\mathbf{a}(\mathbf{x})$ is a non-linear model of parameter vector $\mathbf{x}$, $\mathbf{W}^{-1}$ is a data noise covariance matrix, $\mathbf{x}_a$ is an *a priori* preferred model parameter state vector with attendant covariance matrix $\mathbf{P}^{-1}$, and damping parameter λ. In order to find the minimum, we employ a Gauss-Newton estimator (*Seber and Wild*, 1989) whose k-th iteration is given by

$$\begin{cases} \boldsymbol{\Delta}\mathbf{x}_k &= \left(\mathbf{A}_k^{\mathrm{T}}\mathbf{W}\mathbf{A}_k+\lambda\mathbf{P}\right)^{-1}\left(\mathbf{A}_k^{\mathrm{T}}\mathbf{W}\mathbf{r}_k+\lambda\mathbf{P}\left(\mathbf{x}_a-\mathbf{x}_k\right)\right), \\ \mathbf{x}_{k+1} &= \mathbf{x}_k+\boldsymbol{\Delta}\mathbf{x}_k, \end{cases} \tag{2}$$

where $\mathbf{r}_k{=}\mathbf{d}-\mathbf{a}(\mathbf{x}_k)$ and $\mathbf{A}_k{=}\frac{\partial\mathbf{a}}{\partial\mathbf{x}_k}$ are the residual vector and the Jacobian matrix, respectively, at the nominal state $\mathbf{x}_k$ of the k-th iteration. In gravity field determination, $\mathbf{x}_k$ is the vector with coefficients of a spherical harmonic expansion of the gravity field (and other parameters that are estimated from the data), the matrix $\mathbf{A}_k$ is the matrix with partial derivatives of the measurements with respect to the estimated parameters for iteration k, the matrix $\mathbf{W}$ is the data weight matrix, the vector $\mathbf{r}_k$ contains the data residuals, and $\mathbf{P}$ is the constraint applied in the inversion.

The familiar Kaula constraint is a smoother whose preferred model state is $\mathbf{0}$ with attendant diagonal covariance matrix $\mathbf{K}^{-1}$ such that $K_{ii}=1/\sigma_i^2$, where the σ_i are specified by a power rule. This power rule is based on the observation that the standard deviation of the coefficients of the gravitational potential of spherical harmonic degree n follow the rule B/n^2 with B a constant depending on the planet (*Kaula*, 1966). This constraint has been applied in the estimation of gravity field models of Mercury (*Mazarico et al.*, 2014), Venus (*Konopliv et al.*, 1999), Mars (*Genova et al.*, 2016; *Konopliv et al.*, 2016), the Moon (*Lemoine et al.*, 2013, 2014), and for small bodies (*Konopliv et al.*, 2002, 2014). Assuming that all elements of the M-vector $\mathbf{x}_a$ are non-zero, we consider a modification to this class of constraints whose preferred model states are $\mathbf{0}$ and have attendant inverse-covariance matrices of the following form

$$\mathbf{P} = \mathbf{F}\left(\mathbf{I}-M^{-1}\mathbf{1}\mathbf{1}^{\mathrm{T}}\right)\mathbf{F}, \tag{3}$$

where $\mathbf{F}$ is an $M{\times}M$ diagonal matrix such that $F_{ii}=1/(x_a)_i$, $\mathbf{I}$ is an $M{\times}M$ identity matrix, and $\mathbf{1}$ is an M-vector of ones. It should be clear that $\mathbf{P}$ is a symmetric positive semi-definite matrix of rank $M{-}1$ such that $\mathbf{P}\mathbf{x}_a{=}\mathbf{0}$. This means that it can be factored as $\mathbf{P}{=}\mathbf{G}\mathbf{G}^{\mathrm{T}}$, where $\mathbf{G}$ is an $M{\times}M{-}1$ full-rank matrix whose columns are orthogonal to $\mathbf{x}_a$. The interpretation of $\mathbf{P}$ as an inverse-covariance matrix is understood in the context that in the direction of $\mathbf{x}_a$ the matrix $\mathbf{P}^{-1}$ would assign large variances that tend to infinity in some limiting process. The motivation for the form of $\mathbf{P}$ in eq. 3 will be discussed shortly.

What is of interest now is the behavior of the $\boldsymbol{\Delta}\mathbf{x}_k$ correction term as $\lambda\rightarrow\infty$, that is, as the influence of the constraint increases. For brevity let us define $\mathbf{N}_k=\mathbf{A}_k^{\mathrm{T}}\mathbf{W}\mathbf{A}_k$ and $\mathbf{y}_k=\mathbf{A}_k^{\mathrm{T}}\mathbf{W}\mathbf{r}_k$.

Let us also assume that $\mathbf{N}_k$ is invertible such that it has a Cholesky decomposition of $\mathbf{N}_k = \mathbf{L}_k\mathbf{L}_k^{\mathrm{T}}$, where $\mathbf{L}_k$ is a lower-triangular matrix. Using the Sherman-Morrison-Woodbury formula, we now rewrite $\boldsymbol{\Delta}\mathbf{x}_k$ as

$$\begin{aligned}
\boldsymbol{\Delta}\mathbf{x}_k &= \left(\mathbf{N}_k + \lambda\mathbf{G}\mathbf{G}^{\mathrm{T}}\right)^{-1}\left(\mathbf{y}_k - \lambda\mathbf{G}\mathbf{G}^{\mathrm{T}}\mathbf{x}_k\right), \\
&= \left(\mathbf{N}_k + \lambda\mathbf{G}\mathbf{G}^{\mathrm{T}}\right)^{-1}\left(\mathbf{y}_k + \mathbf{N}_k\mathbf{x}_k\right) - \mathbf{x}_k, \\
&= \left[\mathbf{N}_k^{-1} - \mathbf{N}_k^{-1}\mathbf{G}\left(\lambda^{-1}\mathbf{I} + \mathbf{G}^{\mathrm{T}}\mathbf{N}_k^{-1}\mathbf{G}\right)^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{N}_k^{-1}\right]\left(\mathbf{y}_k + \mathbf{N}_k\mathbf{x}_k\right) - \mathbf{x}_k, \\
&= \mathbf{L}_k^{-\mathrm{T}}\left[\mathbf{I} - \mathbf{L}_k^{-1}\mathbf{G}\left(\lambda^{-1}\mathbf{I} + \mathbf{G}^{\mathrm{T}}\mathbf{L}_k^{-\mathrm{T}}\mathbf{L}_k^{-1}\mathbf{G}\right)^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{L}_k^{-\mathrm{T}}\right]\mathbf{L}_k^{-1}\left(\mathbf{y}_k + \mathbf{N}_k\mathbf{x}_k\right) - \mathbf{x}_k. \quad (4)
\end{aligned}$$

We now take the limit as $\lambda \to \infty$

$$\boldsymbol{\Delta}\mathbf{x}_k = \mathbf{L}_k^{-\mathrm{T}}\left[\mathbf{I} - \mathbf{L}_k^{-1}\mathbf{G}\left(\mathbf{G}^{\mathrm{T}}\mathbf{L}_k^{-\mathrm{T}}\mathbf{L}_k^{-1}\mathbf{G}\right)^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{L}_k^{-\mathrm{T}}\right]\mathbf{L}_k^{-1}\left(\mathbf{y}_k + \mathbf{N}_k\mathbf{x}_k\right) - \mathbf{x}_k, \quad (5)$$

and note that the quantity in the square brackets is a projection matrix onto the null of the column space of $\mathbf{L}_k^{-1}\mathbf{G}$ or onto the column space of its orthogonal complement, which is just $\mathbf{L}_k^{\mathrm{T}}\mathbf{x}_a$. Thus, we can rewrite this as

$$\begin{aligned}
\boldsymbol{\Delta}\mathbf{x}_k &= \mathbf{L}_k^{-\mathrm{T}}\left[\mathbf{L}_k^{\mathrm{T}}\mathbf{x}_a\left(\mathbf{x}_a^{\mathrm{T}}\mathbf{L}_k\mathbf{L}_k^{\mathrm{T}}\mathbf{x}_a\right)^{-1}\mathbf{x}_a^{\mathrm{T}}\mathbf{L}_k\right]\mathbf{L}_k^{-1}\left(\mathbf{y}_k + \mathbf{N}_k\mathbf{x}_k\right) - \mathbf{x}_k, \\
&= \alpha_k\mathbf{x}_a - \mathbf{x}_k, \quad (6)
\end{aligned}$$

where

$$\begin{aligned}
\alpha_k &= \left(\mathbf{x}_a^{\mathrm{T}}\mathbf{N}_k\mathbf{x}_a\right)^{-1}\mathbf{x}_a^{\mathrm{T}}\left(\mathbf{y}_k + \mathbf{N}_k\mathbf{x}_k\right), \\
&= \left(\mathbf{x}_a^{\mathrm{T}}\mathbf{A}_k^{\mathrm{T}}\mathbf{W}\mathbf{A}_k\mathbf{x}_a\right)^{-1}\mathbf{x}_a^{\mathrm{T}}\mathbf{A}_k^{\mathrm{T}}\mathbf{W}\left(\mathbf{r}_k + \mathbf{A}_k\mathbf{x}_k\right). \quad (7)
\end{aligned}$$

In the case of a linear model, where $\mathbf{a}(\mathbf{x})$=$\mathbf{A}\mathbf{x}$, this reduces to

$$\alpha_k = \left(\mathbf{x}_a^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A}\mathbf{x}_a\right)^{-1}\mathbf{x}_a^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{d}, \quad (8)$$

and is seen to be the least-squares solution to the system

$$\mathbf{d} = \mathbf{A}\mathbf{x}_a\alpha + \boldsymbol{\nu}, \quad \boldsymbol{\nu} \sim \mathcal{N}\left(\mathbf{0}, \mathbf{W}^{-1}\right), \quad (9)$$

where $\mathbf{x}$ is now constrained to be a scalar multiple of $\mathbf{x}_a$. The final update is then given by

$$\mathbf{x}_{k+1} = \alpha_k\mathbf{x}_a. \quad (10)$$

It is interesting to note that if $\mathbf{P} = \mathbf{F}^2$ (which is the Kaula form), then as $\lambda \to \infty$ we have $\mathbf{x}_{k+1} \to \mathbf{0}$. This is because $\mathbf{P}$ is full-rank. However, in the case presented here, all directions except that of $\mathbf{x}_a$ are suppressed, thus leaving the solution in the form of $\mathbf{x}_{k+1} = \alpha_k\mathbf{x}_a$, where α_k is determined by the data. The solution has a single degree of freedom in the direction of $\mathbf{x}_a$ because of the infinite variance in this direction due to $\mathbf{P}$. Because all other directions are suppressed, the

preferred model state is indeed $\mathbf{0}$, even though the final model state equals $\alpha_k \mathbf{x}_a$. The factor α_k is determined fully by the data and we thus do not consider this a constraint.

The form of the RM1 constraint ultimately reflects the belief that the bulk crustal density, ρ_c, can describe most of the gravity signal we observe (*e.g., Wieczorek and Phillips*, 1998) through the equation

$$\mathbf{x} = \mathbf{x}_a \rho_c = \mathbf{F}^{-1} \mathbf{1} \rho_c, \tag{11}$$

where $\mathbf{x}$ and $\mathbf{x}_a$ are vectors of the observed gravity and gravity-from-topography coefficients, respectively, and $\mathbf{1}$ is the same as in eq. 3. In order to allow a full description of the gravity field we consider deviations from ρ_c by considering deviations from the basis vector $\mathbf{1}$ by introducing an $M \times M-1$ full-rank matrix $\mathbf{U}$ whose columns are orthogonal to $\mathbf{1}$ such that

$$\mathbf{U}^{\mathrm{T}} \mathbf{1} = \mathbf{0}, \quad \mathbf{U}^{\mathrm{T}} \mathbf{U} = \mathbf{I}, \quad \mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{I} - M^{-1} \mathbf{1}\mathbf{1}^{\mathrm{T}}, \tag{12}$$

which gives

$$\mathbf{x} = \mathbf{F}^{-1} \begin{pmatrix} \mathbf{1} & \mathbf{U} \end{pmatrix} \begin{pmatrix} \rho_c \\ \boldsymbol{\delta\rho} \end{pmatrix}, \tag{13}$$

where $\boldsymbol{\delta\rho}$ absorb deviations from the effect of ρ_c. Inverting eq. 13 allows us to obtain, for instance, ρ_c from the observed gravity coefficients

$$\begin{pmatrix} \rho_c \\ \boldsymbol{\delta\rho} \end{pmatrix} = \begin{pmatrix} M^{-1} \mathbf{1}^{\mathrm{T}} \\ \mathbf{U}^{\mathrm{T}} \end{pmatrix} \mathbf{F}\mathbf{x}. \tag{14}$$

The formula for ρ_c in eq. 14 can be shown in summation notation as

$$\rho_c = \frac{1}{M} \sum_{j=1}^{M} \frac{(x)_j}{(x_a)_j}. \tag{15}$$

This is an interesting formula from the standpoint that since $(x_a)_j$ are generally larger for lower degrees, then the weights applied to $(x)_j$ will be larger for higher degrees, i.e., larger degrees will have more influence on ρ_c. Note that when $\lambda \to \infty$, then $\rho_c = \alpha$.

If we denote

$$\boldsymbol{\rho} = \begin{pmatrix} \rho_c \\ \boldsymbol{\delta\rho} \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} M^{-1}\mathbf{1} & \mathbf{U} \end{pmatrix}, \quad \mathbf{E}_{11} = \begin{pmatrix} 0 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}, \tag{16}$$

where $\mathbf{E}_{11}$ is an $M \times M$ matrix, then we can consider minimizing the cost function

$$J(\boldsymbol{\rho}) = \frac{1}{2} \left(\mathbf{d} - \mathbf{a}(\boldsymbol{\rho})\right)^{\mathrm{T}} \mathbf{W} \left(\mathbf{d} - \mathbf{a}(\boldsymbol{\rho})\right) + \frac{1}{2} \lambda \boldsymbol{\delta\rho}^{\mathrm{T}} \boldsymbol{\delta\rho}, \tag{17}$$

$$= \frac{1}{2} \left(\mathbf{d} - \mathbf{a}(\boldsymbol{\rho})\right)^{\mathrm{T}} \mathbf{W} \left(\mathbf{d} - \mathbf{a}(\boldsymbol{\rho})\right) + \frac{1}{2} \lambda \boldsymbol{\rho}^{\mathrm{T}} \mathbf{E}_{11} \boldsymbol{\rho}, \tag{18}$$

where the first term is data misfit and the second is the squared-length of the $\boldsymbol{\delta\rho}$ vector, thus penalizing deviations away from a pure bulk density. Using eqs. 13 and 14, the k-th iteration of

the Gauss-Newton estimation is given by

$$\begin{cases} \boldsymbol{\Delta\rho}_k &= \left(\mathbf{H}^{-1}\mathbf{F}^{-1}\mathbf{A}_k^{\mathrm{T}}\mathbf{W}\mathbf{A}_k\mathbf{F}^{-1}\mathbf{H}^{-\mathrm{T}} + \lambda\mathbf{E}_{11}\right)^{-1}\left(\mathbf{H}^{-1}\mathbf{F}^{-1}\mathbf{A}_k^{\mathrm{T}}\mathbf{W}\mathbf{r}_k - \lambda\mathbf{E}_{11}\boldsymbol{\rho}_k\right), \\ \boldsymbol{\rho}_{k+1} &= \boldsymbol{\rho}_k + \boldsymbol{\Delta\rho}_k, \end{cases} \tag{19}$$

which reduces to eq. 2. Therefore, we can minimize eq. 17 with respect to $\boldsymbol{\rho}$ by running the Gauss-Newton iterations on $\mathbf{x}$ via eq. 2 and transforming using eq. 14. The reason for working with $\mathbf{x}$ rather than $\boldsymbol{\rho}$ is that we have natural access to Normal Equation systems with respect to the former.

2 Rank minus S (RMS) constraint

One can also imagine applying several separate constraints of this form on several sub-sets of parameters. For example, we could create sub-sets per degree n, to have more degrees of freedom (in the limit of $\lambda \to \infty$) in the effective density spectrum. Without loss of generality, we may assume that the parameter sub-sets are contiguous within each of S sets such that $\mathbf{P}$ and $\mathbf{G}$ take on a block-diagonal forms

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}_1 & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{P}_S \end{pmatrix}, \quad \mathbf{G} = \begin{pmatrix} \mathbf{G}_1 & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{G}_S \end{pmatrix}. \tag{20}$$

The derivation proceeds as before except now the orthogonal complement of $\mathbf{L}_k^{-1}\mathbf{G}$ is the $M \times S$ matrix $\mathbf{L}_k^{\mathrm{T}}\mathbf{X}_a$, where

$$\mathbf{X}_a = \begin{pmatrix} \mathbf{x}_{a_1} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{x}_{a_S} \end{pmatrix}, \tag{21}$$

and $\mathbf{x}_a^{\mathrm{T}} = \begin{bmatrix}\mathbf{x}_{a_1}^{\mathrm{T}} \dots \mathbf{x}_{a_S}^{\mathrm{T}}\end{bmatrix}$. This results in a correction

$$\boldsymbol{\Delta}\mathbf{x}_k = \mathbf{X}_a\boldsymbol{\alpha}_k - \mathbf{x}_k, \tag{22}$$

where

$$\begin{aligned} \boldsymbol{\alpha}_k &= \left(\mathbf{X}_a^{\mathrm{T}}\mathbf{N}_k\mathbf{X}_a\right)^{-1}\mathbf{X}_a^{\mathrm{T}}\left(\mathbf{y}_k + \mathbf{N}_k\mathbf{x}_k\right), \\ &= \left(\mathbf{X}_a^{\mathrm{T}}\mathbf{A}_k^{\mathrm{T}}\mathbf{W}\mathbf{A}_k\mathbf{X}_a\right)^{-1}\mathbf{X}_a^{\mathrm{T}}\mathbf{A}_k^{\mathrm{T}}\mathbf{W}\left(\mathbf{r}_k + \mathbf{A}_k\mathbf{x}_k\right). \end{aligned} \tag{23}$$

Therefore, each sub-set of $\mathbf{x}_{k+1}$ is a scalar multiple of the corresponding sub-set of $\mathbf{x}_a$ due to the S degrees of freedom available to the estimator.

Applying this rank-minus-S (RMS) constraint instead of the RM1 constraint implies that rather than estimating one single α_k factor, we estimate multiple scale factors. These constitute $\rho_{\text{estim}}(n) =$

$\alpha_{kn} \cdot \rho_0$ *if* we take each subset of $\mathbf{x}_{k+1}$ to be that subset of gravity coefficients with the same degree n. We have tested solutions using this RMS constraint instead of the RM1 constraint, but found that in general they resulted in effective density spectra similar to those using the RM1 constraint (if λ was not too large), with larger variations. This shows that RM1 solutions also contain the information on laterally varying density (since the RMS constraint is designed to specifically have variations per degree). If λ increases, the variations in especially the higher degrees increase accordingly, in contrast to the behavior of the RM1 constraint which results in a flat effective density spectrum for large values of λ. We attribute this to the extra degrees of freedom (S compared to only 1) in the solution, and the fact that those extra degrees of freedom are distributed with equal weight, where some (those describing $\rho_{\text{estim}}(n)$ for large n) might be poorly determined from the data. Conceivably one could include further weights where those scale factors for $\rho_{\text{estim}}(n)$ for large n are constrained further.

3 Effective density

For the computation of density from gravity and topography we use the same method that was used for GRAIL in *Wieczorek et al.* (2013). We model the estimated gravity $\mathbf{g}_{\text{obs}}$ from our inversions using satellite tracking data by relating it to gravity induced by surface relief $\mathbf{g}_{\text{relief}}$ (*Wieczorek and Phillips*, 1998). If we express both estimated and modeled gravity in spherical harmonics, assume some density ρ_0 (*e.g.,*, a "unit" density of 1000 kg m^{-3} as we used in our analysis) when we compute $\mathbf{g}_{\text{relief}}$, and assume a constant crustal density ρ_{estim}, then for each degree n and order m, observed gravity is related to the gravity induced by the relief following:

$$g_{\text{obs},nm} = \rho_{\text{estim}}\ g_{\text{relief},nm} + I_{nm} \tag{24}$$

where I_{nm} is that part of the estimated gravity not modeled by topography, and it is assumed to be uncorrelated with $\mathbf{g}_{\text{relief}}$ and to have a zero mean. An unbiased estimate for the crustal density at each degree n then follows from *Wieczorek et al.* (2013):

$$\rho_m(n) = S_{g_{\text{obs}}\ g_{\text{relief}}}(n) / S_{g_{\text{relief}}\ g_{\text{relief}}}(n) \tag{25}$$

where crosspower $S_{ab}(n)$ (or autopower, if both quantities a and b are the same) is defined as:

$$S_{ab}(n) = \sum_{m=-n}^{m=l} a_{nm} b_{nm} \tag{26}$$

With equation (24), and the fact that for $\mathbf{x}_a$ we chose $\mathbf{g}_{\text{relief}}$, it now also follows in a straightforward way that the factor α_k determines the actual model density ρ_{estim} through $\rho_{\text{estim}} = \alpha_k \cdot \rho_0$, and this is how we obtain our estimates for the effective density spectrum.

4 Obtaining the bulk density

The average bulk crustal density for the Moon was derived by averaging the GRAIL-derived effective density spectrum over the degree range n = 150-310 (*Wieczorek et al.*, 2013). Global admittance values for Mars become stable after around degree n = 25 (*e.g., Konopliv et al.*, 2006), and we find a value for the lunar density from our SGM150J-RM1 model over the range n = 60–130 that is very close to the GRAIL value, indicating that we can use similar degree ranges for our Mars models (with have a maximum degree of either 120 or 150). Averaging the density spectrum over the degree range n = 30–115 for each of our clone models, and taking the average of that set, results in an average crustal density of 2536 kg m^{-3}. Following the GRAIL results, the associated error is computed from the density variations over the same range around this average value, resulting in an error of 167 kg m^{-3}.

This is a uniformly weighted average, the same as was done for the GRAIL results. As Fig. 2A in the main text shows however, the variations are larger than those for the Moon, and, because of noise in the data, they increase for $n \geq 110$. Alternatively, we can therefore also compute an average weighted by an error estimate (or rather, the inverse thereof). We obtain error estimates for the density at each degree from our set of clones: for each clone, we compute the effective density spectrum, and then we compute the standard deviation at each degree around the average of all clones. Figure S12 shows the Mars effective density spectrum together with such an error estimate. We can then again compute the average over the same degree range n = 30–115, weighted by the inverse of the error at each degree. This results in a higher bulk density of 2660 $\pm$ 209 kg m^{-3}, where the error is again obtained from the variations around the average value. Figure S13 shows an alternative error spectrum, where the errors are taken as the minimum and maximum value for each degree that occur in the set of clone models. This produces a conservative error spectrum, but the weighted average is not much different, at 2657 $\pm$ 206 kg m^{-3}, probably because mostly this leads to scaled errors (for example at 3 or 4 times the standard deviation). However, the errors for the coefficients of gravity models increase with increasing degree due to decreased data sensitivity, see for example Fig. 5 in *Genova et al.* (2016), and this means that such a weighted average will always obtain results where higher degrees have less influence. Apart from the consideration of errors in the models however there is not necessarily a strong reason to take such a weighted average, because at higher degrees it is expected that the model describes shallower parts of the crust, and there is thus no reason to expect them to contribute less to the average crustal density.

As explained in section 1, our choice of constraint also provides a different way of computing the bulk density, as given by eq. 15. As stated, this gives more influence to the higher degrees. Computing the average density in this way for all the clones, again using the degree range n=30–115, produces an average density of 2582 kg m^{-3}, which is in between the uniform and weighted average values. The spread around this value in the clones set is very small, at only 9 kg m^{-3}. We adopt this value as our preferred bulk crustal density. We use the larger error value to reflect current uncertainties in the value due to noise and the variations in density that are inherent to the models.

The range of degrees over which the averages, or over which eq. 15, are taken do influence the

obtained bulk density to some extent: due to the downward trend, an average obtained from starting at a higher degree will be lower. Likewise, when the gravity field of Mars is determined up to resolutions similar to GRAIL's at the Moon, it is likely that the average density will be lower, with even lower surface densities, as was the case at the Moon (*e.g., Besserer et al.*, 2014). On the other hand, our estimate can be biased upwards by the densities of the volcanoes. The density map shown in Fig. 2C shows lower densities in other areas. Our estimates however do not change within the given error bounds when different degree ranges are used, given the current knowledge of the gravity field of Mars (where we cannot expect GRAIL-like resolutions). Moreover, the asymptotic density that we find, at 2597 kg m^{-3}, is close to our average.

5 GRAIL primary mission data only and the RM1 constraint

To further verify that our constraint estimates the correct crustal density, in addition to applying it to a pre-GRAIL least-squares system, we also apply it to the system of the GRAIL primary mission model GRGM660PRIM (*Lemoine et al.*, 2013). We again compare the results of this analysis with those from the model based on GRAIL extended mission data, GRGM900C (*Lemoine et al.*, 2014). This model far surpasses the resolution of the primary mission models, and it serves as truth, as this extended mission model is mostly constraint-free for the degree range of GRGM660PRIM. In Fig. S1 we show the effective density for GRGM660PRIM, GRGM900C, and a model based on the GRGM660PRIM system that uses our RM1 constraint. The model GRGM660PRIM was derived using a Kaula constraint applied for $n > 330$, and Fig. S1 clearly shows that the effective density drops for high degrees. The model that uses our RM1 constraint (also applied for $n > 330$) however follows the extended mission model result much closer, over its full degree range. The results in Fig. S1 used a damping factor $\lambda = 1$. We use a slightly lower value for λ in this case, compared to the value $\lambda = 10$ for the SGM150J case, because the GRAIL system has much better resolution than the SGM150J system. We thus increase the influence of the constraint for the SGM150J case by increasing λ slightly. Using $\lambda = 1$ for the SGM150J case results in slightly more variations in the density curve.

6 The RM1 constraint and parameter dependencies

These two different lunar cases (for SGM150J and GRGM660PRIM) show that our RM1 constraint gives a reliable estimate of the effective density, as was also exemplified by the quoted values for the density from the SGM150J RM1 results in the main text. In addition, we stress that our RM1 constraint does not bias the results to the *a priori* density value used for computing gravity-from-topography: choosing a density different from 1000 kg m^{-3} in our computation of gravity-from-topography is equivalent to using a slightly different damping factor λ, but the same effective density is still obtained. For large damping factors, the effective density becomes flat by definition (it is given by α_k times the used "unit" density, at all degrees n), and the density value we then

obtain for the GRGM660PRIM results is close to the value of the density at the degree where the constraint is applied (Fig. S2). Because at very low degrees gravity is controlled by large scale signals in the deep interior and thus does not necessarily correlate well with topography, the RM1 constraint should not be applied over the full degree range of the model (*i.e.* for $n \geq 2$).

For the Mars results we applied the RM1 constraint from different degrees (*e.g.,* $n \geq 50$ and $n \geq 85$) and this results in effective density spectra that are very close to one another, as can be seen in Fig. S4A. Correlations for both these solutions are shown in Fig. S4B, and it can be seen that for the solution with RM1 applied for $n \geq 85$ the correlations jump up, whereas the correlations gradually increase for the solution with RM1 applied $n \geq 50$. The effective density estimate however is not affected at all by this jump as shown in Fig. S4A.

The damping parameter λ is discussed throughout the main text and this document. The factor determines the relative weight between the data system and the constraint system. As discussed extensively, as λ goes to infinity, the correlations between the solution $\mathbf{x}$ and the solution $\mathbf{x}_a$ go to ± 1, but the path of how the correlations change depends on the relative contributions of the data and constraint. In Figure S3 we show the effective density spectrum and correlations for the SGM150J case. We applied different values of λ to obtain the models. From this Figure it can be clearly seen that the correlations gradually become unity, and that they only become unity for large values of λ. For smaller values (1, 10 and 100) the correlations take on the shape of those provided by GRGM900C, with the correlations for $\lambda = 10$ falling on top of them. We note that the correlations are not constrained to follow the GRAIL shape. Moreover, the effective density spectrum is not much affected by the different correlations. Variations within the effective density per degree, which are interpreted as indicating lateral density variations (see also below, Section 8), become smaller as λ increases, but they are still substantial at $\lambda = 10$, larger than those from the "truth" model GRGM900C. The variations exhibited in the density spectrum for the RM1 model thus are a combination of actual lateral variations, and the influence of noise on the data. Because the variations only decrease for very large values of λ and because we prefer smaller values ($\lambda = 1$ for the Mars results discussed later, $\lambda = 10$ for the SGM150J), it is not immediately clear that our RM1 models would suppress the lateral variations. It is likely that in the case of Mars, the actual variations in the effective density spectrum will become smaller for higher degrees, just like it is the case for the GRAIL results, but we will not know this until we have GRAIL-like resolution data at Mars. The average values of the density are not influenced very much by these variations, as our averaged SGM150J results show, but of course our error estimate is.

7 Kaula and the RM1 constraint

As mentioned in the main text and here in section 1, the Kaula constraint that is generally applied in the determination of gravity field models from satellite tracking data forces the solution towards zero, because the preferred model state is $\mathbf{0}$. We illustrate this further by comparing the effective density spectrum of a Kaula constrained model with that of an unconstrained model in Fig. S5A. The unconstrained GMM-3 model was obtained by inverting the data matrix, which is

possible owing to the global coverage of the Mars tracking data, even if the southern hemisphere has better coverage with lower altitude data than the northern hemisphere (*Genova et al.*, 2016). The effective density for the unconstrained GMM-3 model shows large variations, yet its effective density spectrum is extended to higher degrees than that for the standard Kaula constrained model GMM-3. This clearly shows that the Kaula constraint suppresses the information in the higher degrees, as the density spectrum quickly decreases for GMM-3. The density spectrum for the RM1 constrained model reduces the variations in the density values per degree (as also demonstrated in Fig. S2), and it extends the density spectrum further. This indicates that the RM1 constraint indeed extracts information from the data, because the data-only system of the unconstrained GMM-3 model indicates similar densities although with larger (unrealistic) variations. The RM1 constraint thus allows us to extract the density values by looking at the part of the model space that aligns with gravity-from-topography (our choice for $\mathbf{x}_a$ from section 1). The determined density is furthermore not dependent on the density ρ_0 that was used when generating $\mathbf{x}_a$ because the factor α_k is determined freely from the data, as explained in the main text and in section 1.

For degrees smaller than about 85, all models produce the same density values. The Kaula constraint, applied for degrees $n \geq 90$ for GMM-3 (*Genova et al.*, 2016), does not have much of an influence in this range if the correlations between higher degree coefficients are not too strong (*i.e.,* if they are separable). The RM1 constraint when applied for $n \geq 50$ can have an influence in this range, depending on the constraint factor λ. A small enough factor will still allow the solution to be mostly determined by the data. As we increase lambda, we will obtain a flatter and flatter density spectrum, as we showed in Fig. S2. However, we do not expect the models to be different in this degree range, because the data themselves are strong enough there to fully determine the solution. Indeed, the GMM-3 model itself thus already produces a good estimate of the bulk crustal density. Our RM1 constraint in this case helps to extend the degree range over which we can average the results, making the estimate more robust since it is more likely that gravity and topography are correlated better at higher degrees. In addition, our RM1 constraint helps to determine regional variations (see the main text and below), something we can not do with confidence with the Kaula constrained model. To illustrate this further, we also included results for an extended model in Fig. S5A. We processed the Mars Reconnaissance Orbiter (MRO) data in the same way as was done for GMM-3, but instead generated partial derivatives of the measurements with respect to the coefficients of a spherical harmonic expansion of the gravity field up to a maximum degree and order of 150 instead of 120. We did this only for MRO because it has the lowest altitude of the spacecraft data used in GMM-3. We then again applied the same RM1 constraint, and Fig. S5A shows that the effective density spectrum is further extended. Our density results however do not change significantly when we use this degree and order 150 model over the results of the 120 model. The correlations with topography, shown in Fig. S5B, are also further extended for this 150 model. We note that the correlations for the Kaula constrained and the unconstrained model are nearly identical, despite different density spectra.

8 Local density estimates

While the RM1 constraint depends on gravity-from-topography that assumes a constant density, our results can still derive local variations. As we have shown, for the case $\lambda \to \infty$ the density spectrum collapses, and we find indeed a constant crustal density. Decreasing the value for λ allows variations in the effective density spectrum, which come from lateral variations. We keep in mind that some of the variations can be through noise, or the effects of bottom-loading, but these variations are also present in the lunar spectrum, and for Mars they are present in the degree range that is fully determined by the data ($n \leq 80$). Localization can then extract the local density variations by comparing the estimated gravity field with the gravity-from-topography field (using a "unit" density), in principle determining the scale factor for that chosen density at each location. The same localization approach, with the same assumption of a constant density when computing gravity-from-topography, was used for GRAIL (*e.g., Besserer et al.*, 2014; *Han et al.*, 2014).

To compute regional values, as we do for the Tharsis complex for Mars (Fig. S10), we apply localized spectral analysis (*Wieczorek and Simons*, 2005). Similar to the method of computing the localized admittance between gravity and topography, we compute the ratio as given in eq. 25 with each function multiplied by a spherical cap localization window at a given location using a given radius for the cap, using the freely available software SHTOOLS (*Wieczorek and Meschede*, 2015). We used two different approaches using this windowing method: one to compute the spectrum for Mons Pavonis as given in Fig. S10, and one to compute the laterally varying density map as given in the main text in Fig. 2C.

For the Tharsis region, we apply one taper. We choose a cap centered on Pavonis Mons (247.04°E, 1.47°N) with a spherical cap radius of 20°, and the spherical harmonic bandwidth of the localizing window L_{win} is 20, resulting in a concentration factor larger than 99.99% for the first (and only) taper. We obtain our density estimate for the Pavonis area by averaging the effective density spectrum between degrees $n = 50 - 85$ (uniformly weighted, so each degree contributes in the same way). To construct a global map in a similar way, we repeat this localization procedure on nodes spaced at 5° intervals, covering the whole globe. We use the same windowing size L_{win}=20, again with a cap radius of 20°. For plotting purposes, we resample the output grid from a resolution of 5° to 0.5°. Using one taper only in this procedure, we obtain wildly varying densities, due to varying correlations over the areas and effective density spectra that are not well-behaved. By averaging over multiple tapers instead of one (we settled on 30 tapers), the correlations and effective density spectra are in general better behaved. We again average the effective density spectrum between $n = 50 - 85$. However, the resolution for the Mars models is limited, making it difficult to robustly perform a multi-taper analysis (*e.g., Wieczorek*, 2008). The multi-taper result still shows varying effective density spectra in many locations, resulting in many areas with relatively low densities. Although a recent analysis of the Medusae Fossae Formation produces a density of around 1800 kg/m^3 (*Ojha et al.*, 2017), attributing this to a pyroclastic deposit (dense material with high porosity), our map shows widespread low densities. While the result is broadly consistent with expectations, we stress that the current model resolution does not allow a similar robust treatment as was possible with the GRAIL models (*Besserer et al.*, 2014), and our density map thus mainly

serves as an example of how density may vary laterally on Mars. Correlations between gravity and gravity-from-topography are varying spatially as well (see Fig. 2D in the main text), complicating especially the interpretation of gravity in the northern highlands.

Considering only those locations that have average correlations in a given degree range larger than 0.8 (a signal-to-noise ratio better than 1.78 (*Wieczorek*, 2008; *Grott and Wieczorek*, 2012)), and that have variations in correlations in that degree range that are smaller than 0.05 in a root-mean-square sense, we obtain a density map (shown in Fig. S11) that leaves the northern hemisphere as well as parts of the southern hemisphere undetermined. This map shows the regions on Mars where we are more confident about our density estimates. Not surprisingly, it covers mostly the volcanic regions, and some of their surroundings. According to this map, densities in the northern hemisphere can not be determined with the same amount of confidence. If we increase the threshold for correlations to 0.9 (not shown), the area shrinks further. However, we used our fully global laterally varying density map in one of our crustal thickness estimates in Fig. 3 in the main text because the densities are broadly consistent with geophysical and geological expectations, and to show how an example of a varying density map affects the crustal thickness estimates.

9 Exploring models of porosity stratification

A decrease in density with increasing spherical harmonic degree is an indication that porosity plays an important role in determining the crustal structure (*Wieczorek et al.*, 2013; *Besserer et al.*, 2014; *Han et al.*, 2014). Here, we explore the analytic models from *Han et al.* (2014) for the effective density to determine whether we can estimate parameters related to the increase of density with depth, since our effective density spectrum shows a decrease with increasing degrees (see Fig. 2A in the main text). Two straightforward models can be readily used to describe the dependence of density with depth: a linear model and an exponential one. The linear model assumes a constant increase in density with depth (*Besserer et al.*, 2014):

$$\rho(z) = \rho_{\text{surf}} + az \tag{27}$$

where ρ_{surf} is the surface density, a the gradient, and z the depth (positive for increasing depth). The exponential model describes an asymptotic exponential increase of density with depth (*Besserer et al.*, 2014):

$$\rho(z) = \rho_{\text{surf}} + (\rho_0 - \rho_{\text{surf}})(1 - e^{-z/d}) \tag{28}$$

where ρ_0 is the density at depth (the zero-porosity density) and d is the e-folding scale. The latter model especially can be readily related to a model of pore closure due to compaction (*Han et al.*, 2014). Analytic models for the effective density spectrum for both can be readily derived, valid either globally (*Han et al.*, 2014) or locally in a more Cartesian approach (*Besserer et al.*, 2014).

Ideally the determination of parameters such as ρ_{surf} and a or d would be done in a localized fashion similar as used to derive the GRAIL results (*Besserer et al.*, 2014; *Han et al.*, 2014) and as

how we derived our laterally varying density map, since it is to be expected that these parameters vary spatially. Volcanic complexes likely have different depth-dependencies than the highlands. However, while we were able to derive a density map that is broadly consistent with the expectations, see also section 8, we found that the current resolution of the Mars models is not sufficient to reliably determine multi-parameter depth-dependencies. We illustrate this with a misfit plot in Fig. S14 where we show the misfit between analytic models and the effective density spectrum for our RM1 model, taken between degrees $n = 60$ and $n = 120$, for an exponential depth-dependency model (assuming $\rho_0 = 2900$ kg m^{-3} in this case). While a minimum exists, it is not well-defined, and there are many models with widely varying values for ρ_{surf} and d that fit the data equally well. Misfit plots for the linear model are similar. The models we find are consistent with increasing densities with depth, and in general we find smaller e-folding scales for Mars than for the Moon (which suggests that density for Mars increases faster than at the Moon, which is to be expected considering likely higher pressures and temperatures, and higher surface gravity). However, we applied the same modeling to a GRAIL model using the same degree range ($n = 60 - 120$) and found that while the misfit for the Moon is much narrower, we could not reproduce the published results (*Besserer et al.*, 2014) in this degree range. We thus find we can not estimate the parameters of a depth-dependent model with confidence. Our results indicate the likelihood that density increases with depth, but for our analysis of for example crustal thickness, we assume crustal densities that remain constant with depth. Future missions that increase the current resolution of gravity field models of Mars can resolve the issue of stratification within the crust in a similar way as GRAIL did for the Moon.

10 Crustal thickness modeling

We use the methodology of potential anomalies on the sphere due to finite amplitude relief and downward continuation of the Bouguer anomaly (the estimated free-air anomaly minus the contribution in gravity of the relief) to determine the thickness of the crust, following *Wieczorek and Phillips* (1998). We did not filter the anomalies, but instead use gravity up to degree $n = 90$, and determine the expansion of the sub-surface relief that determines the crustal thickness up to degree n=80. We use a mantle density of 3500 kg m^{-3} (*e.g., Neumann et al.*, 2004). For crustal thickness modeling one also needs to assume either a mean crustal thickness, or a certain crustal thickness at a given location (*Wieczorek*, 2015). We select the average crustal thickness to be 42 km, following the results for GMM-3 as given by *Genova et al.* (2016). We note that these are choices in the process to determine the crustal thickness, and these choices are not constrained by our data.

We present three different crustal thickness models in Fig. 3 in the main text: one based on the standard crustal density of 2900 kg m^{-3}, one based on a density map that uses the new lower value of 2582 kg m^{-3} everywhere except in designated volcanic areas where a value of 2900 kg m^{-3} was used (this value, lower than 3200 kg m^{-3}, was chosen to prevent convergence issues; see Fig. S15 for the density map), and one based on the laterally varying density map we derived from our RM1 solution (again, capping the density values at 2900 kg/m^3, for comparison with the

other density models). The use of laterally varying density maps has been incorporated into the crustal thickness computations (*Wieczorek and Phillips*, 1998; *Wieczorek*, 2015). Using the lower constant crustal density for Mars of 2582 kg m^{-3} for the whole crust reveals a mismodeling of the crustal roots beneath volcanic complexes, as the resulting thickness would appear to be too low for the Tharsis complex on Mars for instance: the topography of the volcano should at least be partially compensated by thicker crust with a density smaller than that of the mantle. Our localized density estimates (both the one over Pavonis, and the example from the analysis of laterally varying densities) confirm this departure from the low average crustal density for the volcanoes, and we thus used laterally varying density maps that take this into account.

11 Supplementary Figures

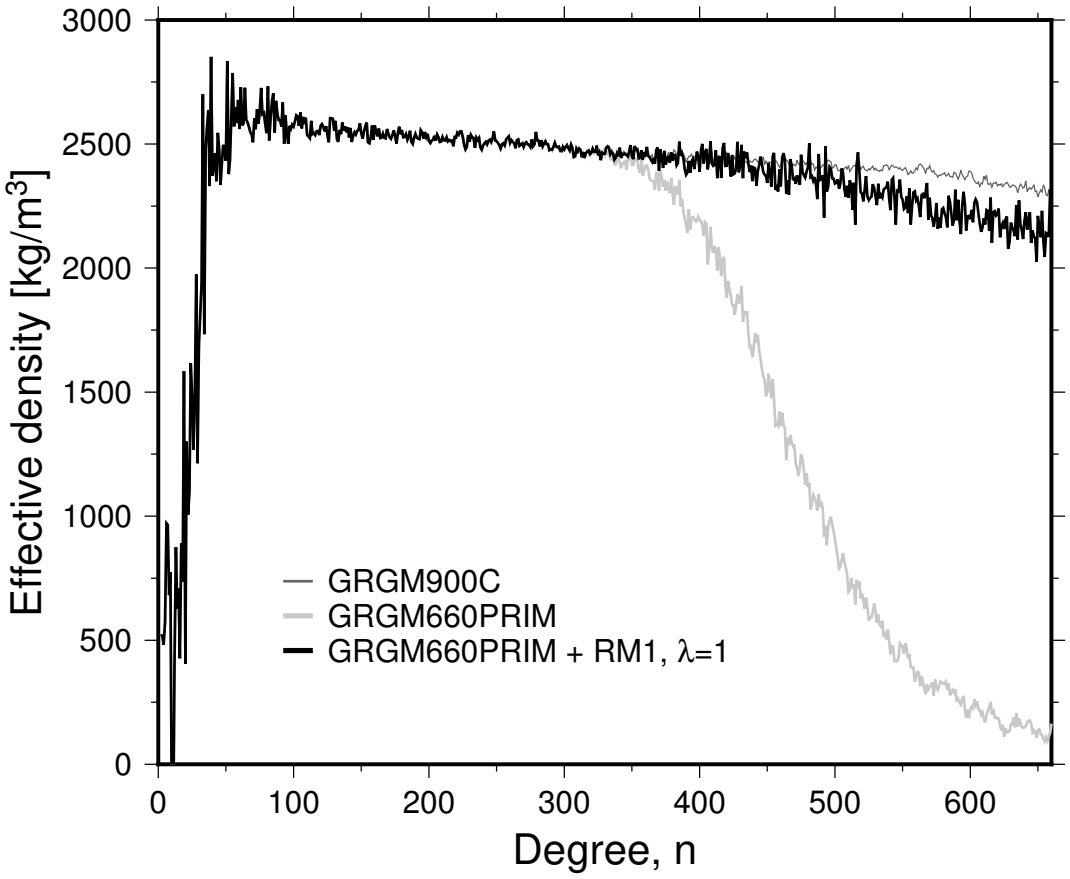

Figure S1: Effective density for various lunar gravity field models: a GRAIL model using only primary mission data (GRGM660PRIM), a GRAIL model with extended mission data (GRGM900C), and a model based on the GRGM660PRIM system using our new "rank minus one" (RM1) constraint.

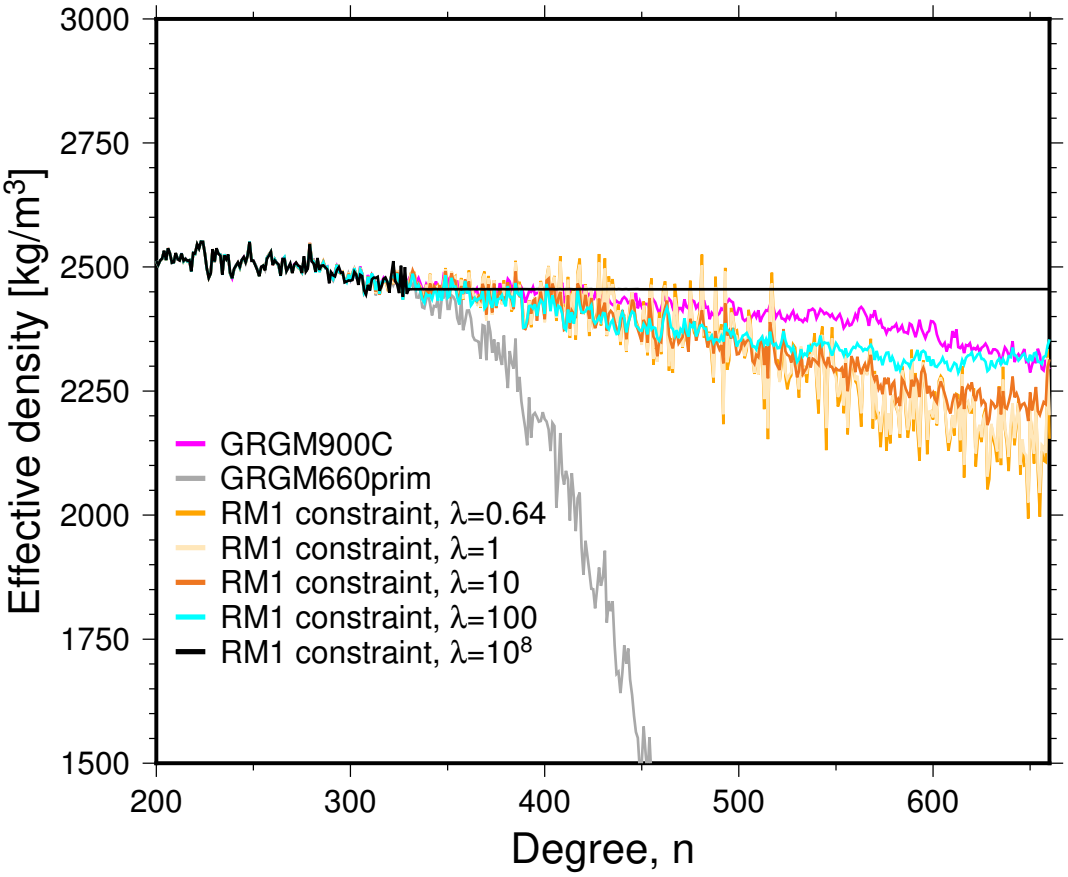

Figure S2: Effective density for the GRGM660PRIM model with the RM1 constraint, applying various damping factors λ. With increasing λ, the effective density per degree n becomes constant.

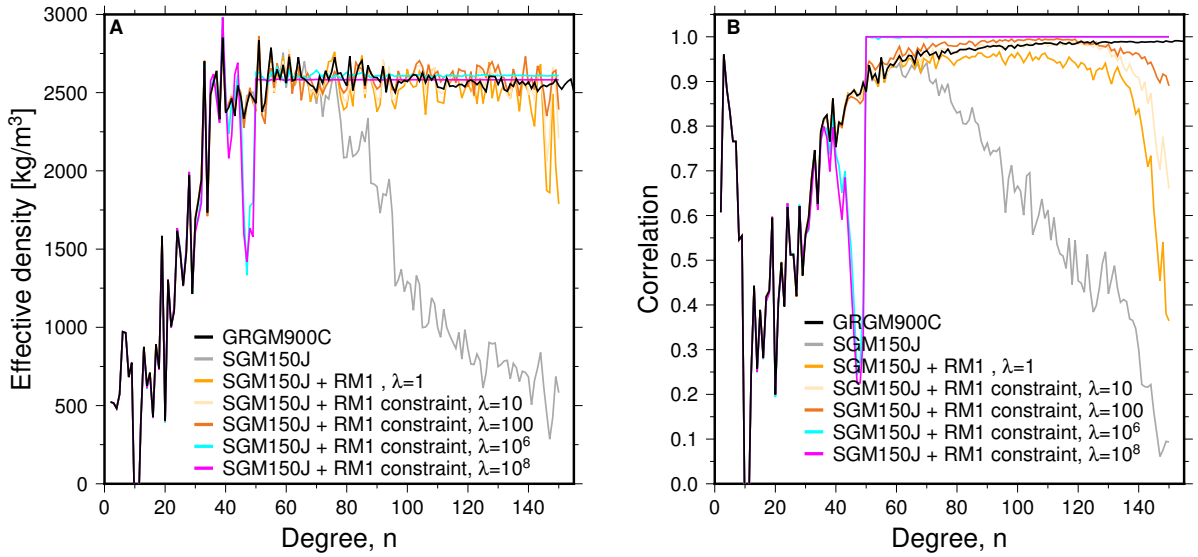

Figure S3: (**A**) Effective density for the SGM150J results for various damping factors λ. (**B**) Correlations between gravity and gravity-from-topography for the same models.

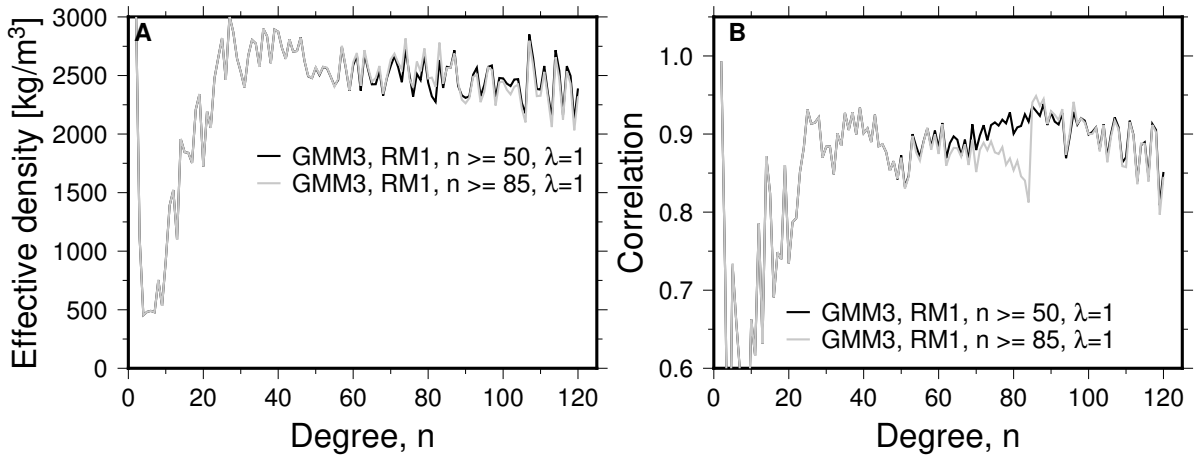

Figure S4: (**A**) Effective density for the Mars results for models with the RM1 constraint applied for either $n \geq 50$ or $n \geq 85$. (**B**) Correlations between gravity and gravity-from-topography for the same models.

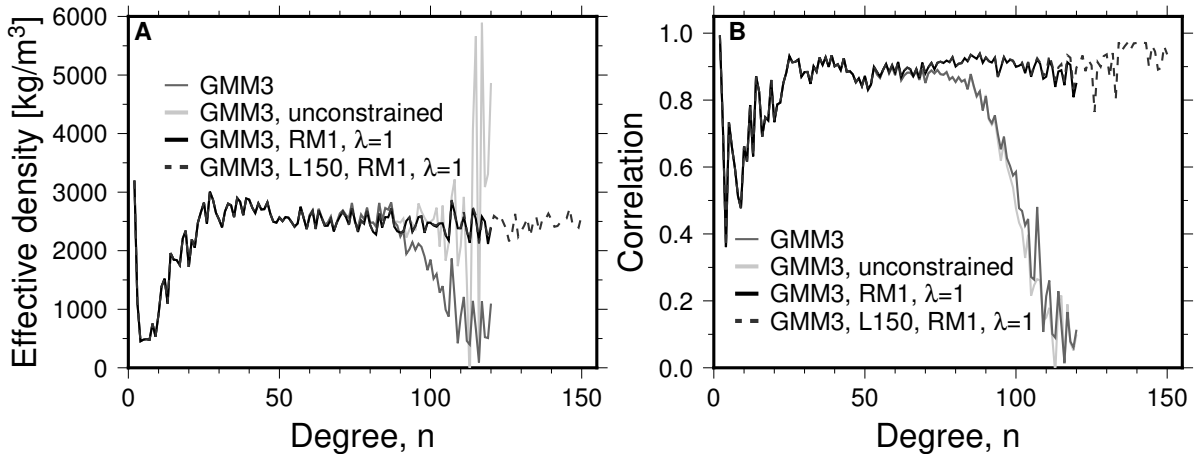

Figure S5: (**A**) Effective density for the Mars results models with a Kaula constraint (GMM-3), the RM1 constraint, or no constraint. The "L150" model is an extended GMM-3 model up to degree and order 150. (**B**) Correlations between gravity and gravity-from-topography for the same models.

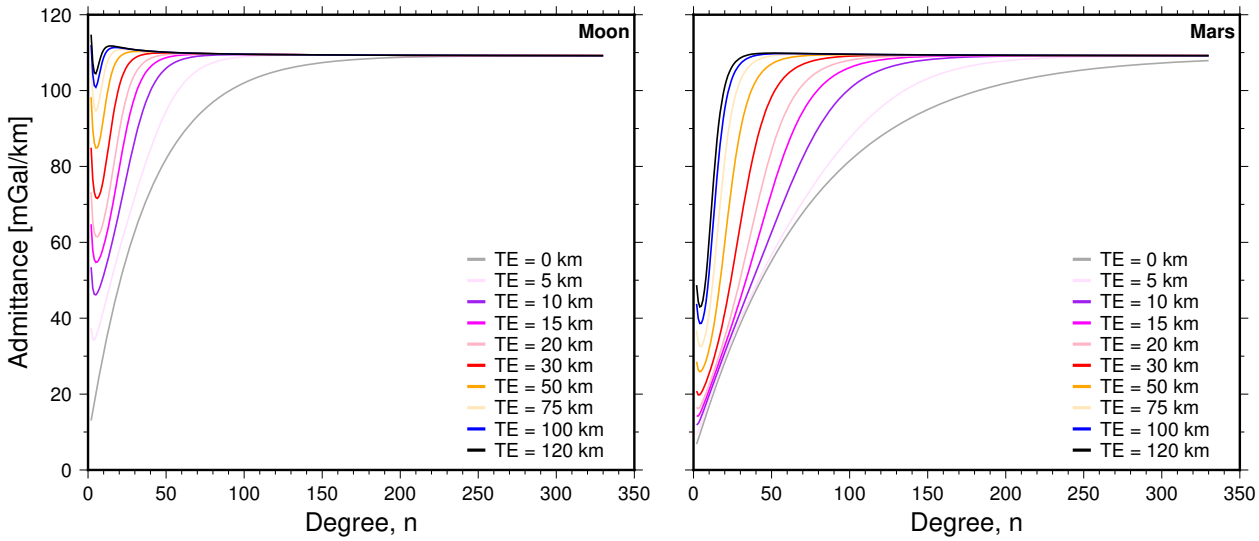

Figure S6: Theoretical admittance for a **Moon** and a **Mars** case, using various values for the elastic thickness. Crustal thickness was set to 45 km, and crustal density to 2600 kg m^{-3}. For Mars, the asymptotic value of the admittance is reached for a larger degree n.

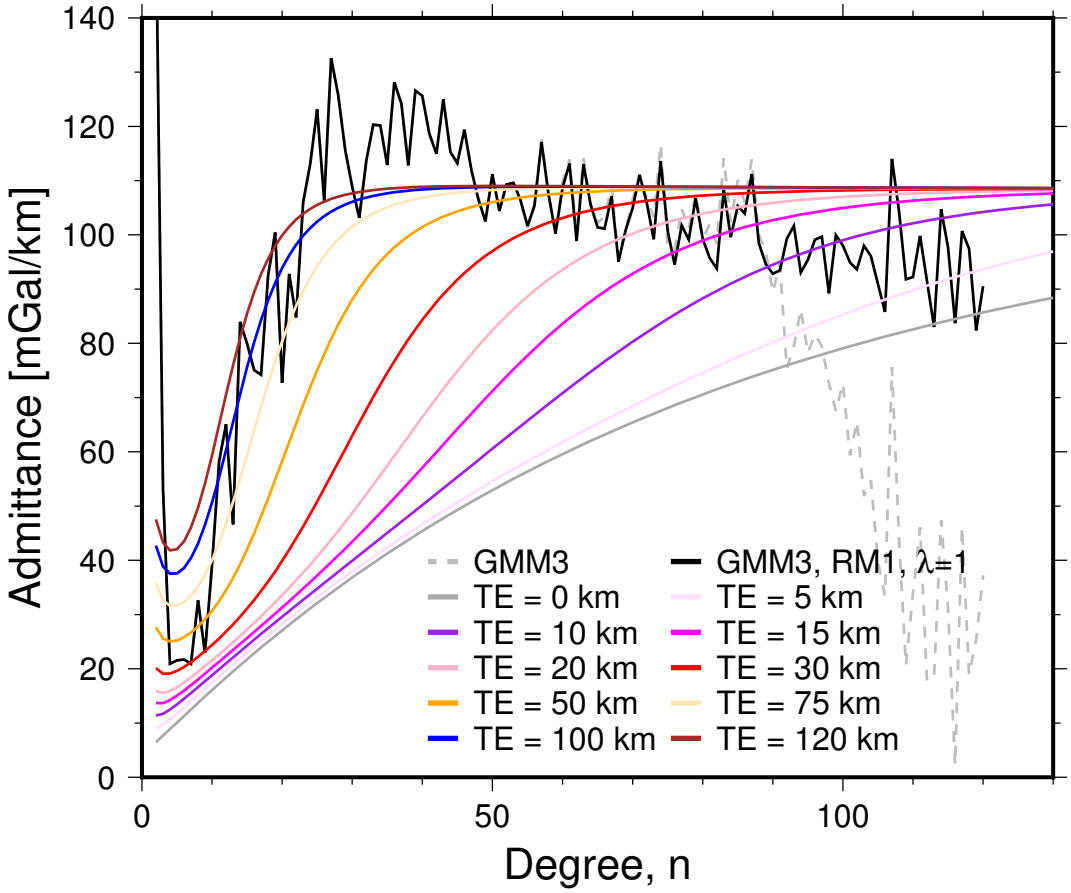

Figure S7: Admittance for Mars gravity models GMM-3 and our new RM1 model, together with theoretical admittance using a crustal thickness of 43 km and a crustal density of 2580 kg m^{-3} for various values of elastic thickness T_e. The measured admittance is best represented by models with high T_e.

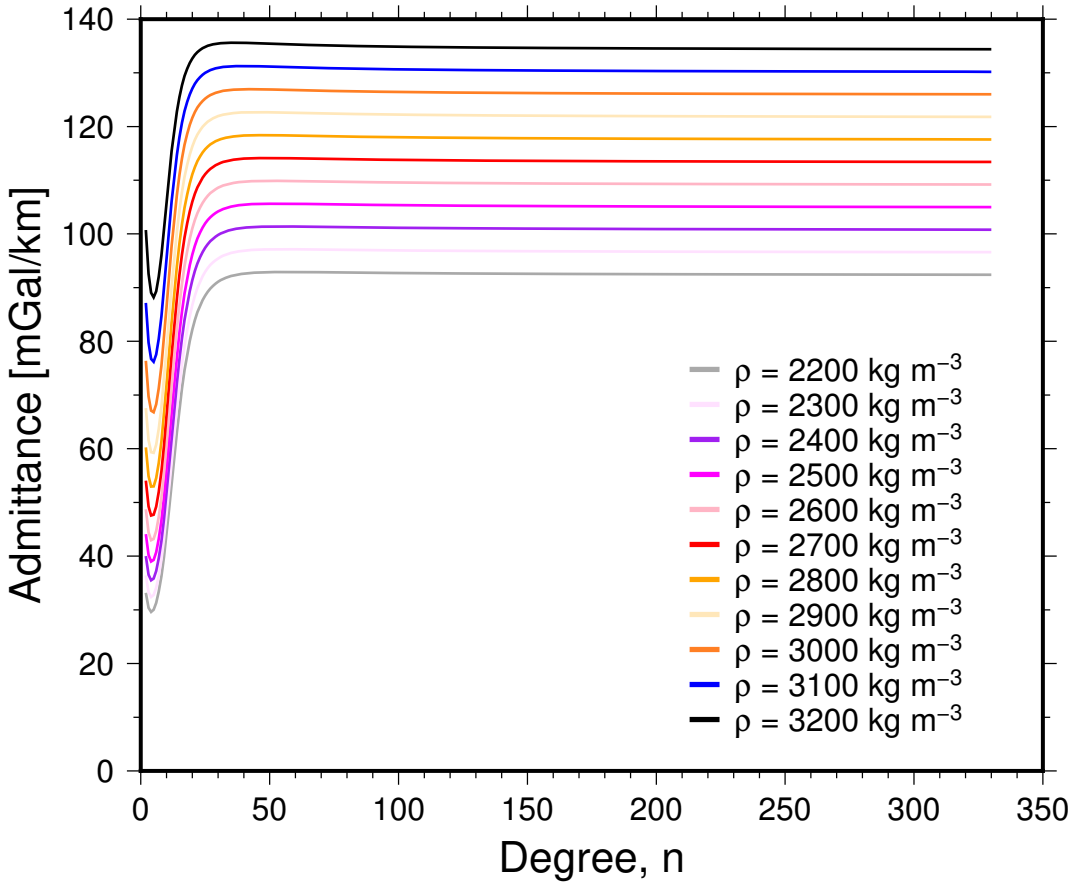

Figure S8: Theoretical admittance for various densities. Elastic thickness was set to 120 km for this plot, and crustal thickness was set to 45 km. The asymptotic value of the admittance only depends on the crustal density. The asymptotic values are the same for different crustal thickness values. Our RM1 constraint estimates this asymptotic value and thus circumvents dependence on elastic thickness.

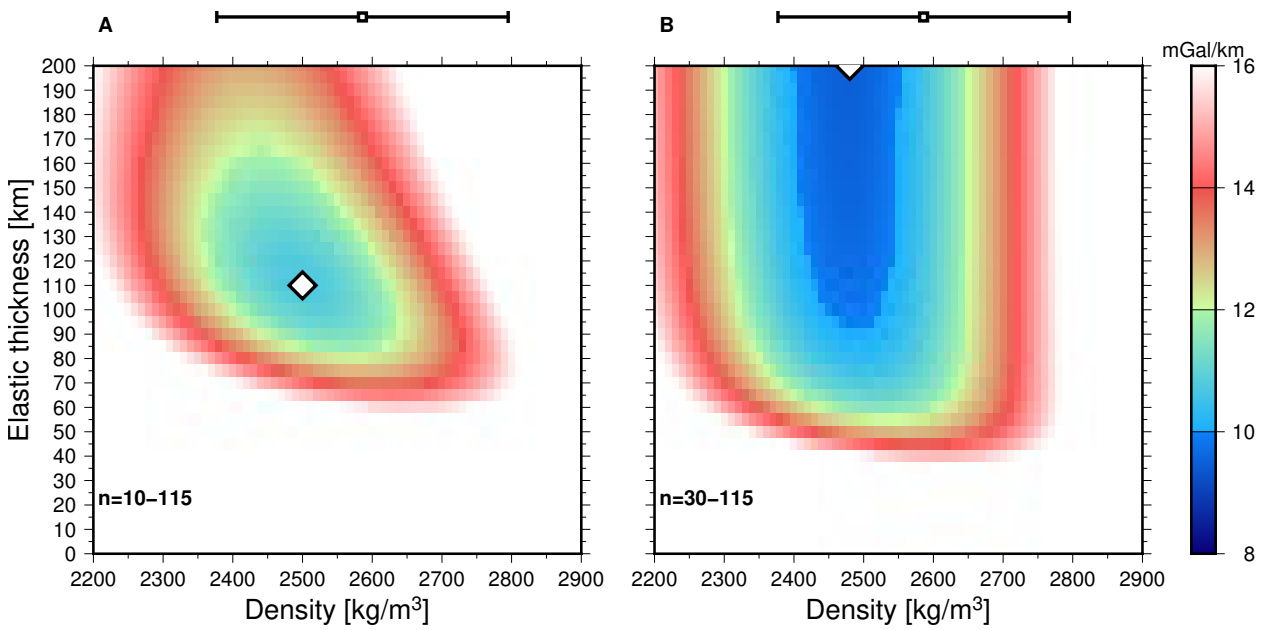

Figure S9: Misfit between measured admittance using the RM1 model and theoretical models using a crustal thickness of 43 km and varying values for the crustal density and elastic thickness T_e. The lowest misfit location is indicated with a diamond. (**A**) Results for the misfit between the admittances using the degree range n = 10–15. (**B**) The same but for the range n = 30–115. Our density estimate and its associated error using the effective density spectrum are indicated above each map (square with error bars). Both degree ranges exclude small values of T_e and estimated densities are in agreement with the estimate from the effective density spectrum. The lowest misfit in (**B**) occurs for the highest T_e value in the search space used, indicating the solutions are not sensitive to T_e for that range.

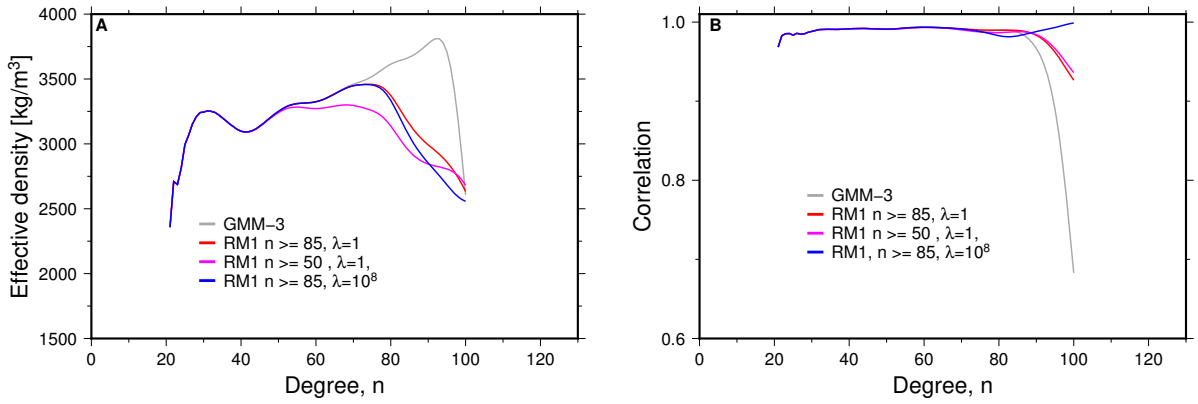

Figure S10: (**A**) Effective density over the Pavonis area. (**B**) Correlations with gravity-from-topography. Both were computed from localized spectra (*Wieczorek and Simons*, 2005) using a cap radius of 20° and a spherical harmonic bandwidth of the localizing window $L_{\rm win}$ = 20. While GMM-3 already gives a reasonable estimate for the lower degrees, the quick drop in correlations with gravity-from-topography results in abnormally high densities. The results for the RM1 constraint are much more stable. The estimated density of 3231 $\pm$ 95 kg m^{-3} is in agreement with other estimates in the literature. We obtain our estimate for the RM1 model that applies the constraint for $n \geq 50$, by taking the average effective density and its variation around this average in the degree range $n = 50 - 85$. Using different degree ranges does not affect the density within the listed error.

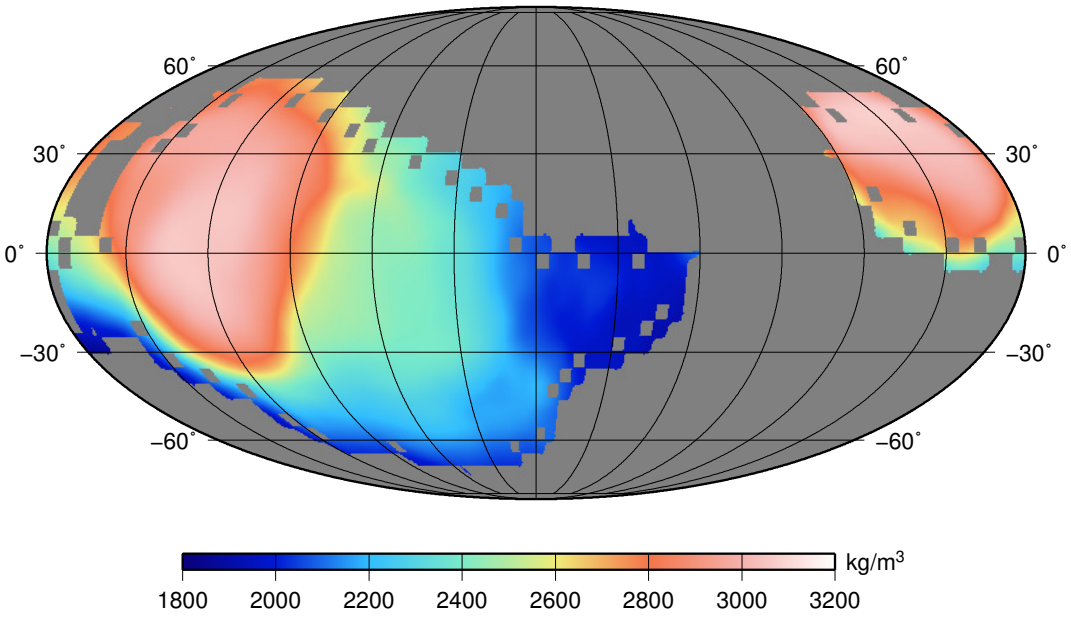

Figure S11: Laterally varying densities (using our RM1 model from the degree and order 150 expansion), the same as in Fig. 2C in the main text, but in this case we retain only those locations where averaged correlations with topography between degrees $n = 50$ and $n = 85$ are larger than 0.8 and variations between the same degrees are smaller than 0.05. This map thus shows the regions where we are more confident about our local density estimates.

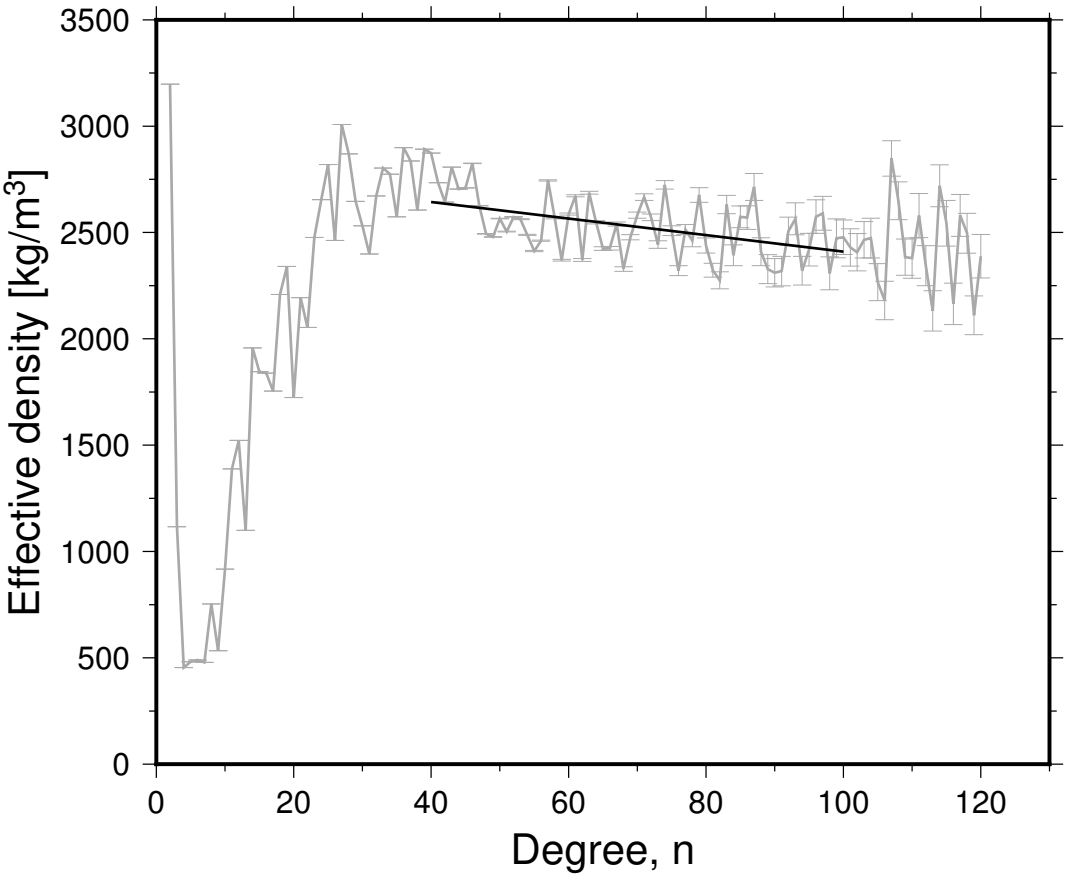

Figure S12: Effective density spectrum for Mars from Fig. 2A in the main text, with the errors computed as standard deviation in the effective density values per degree from the full set of clone models. The average of the trend in the set of clones, fitted between degrees n=40–100, is also indicated, resulting in a function that is described by $\bar{\rho}(n) = 2836 - 4.18 \cdot n$ [kg m^{-3}].

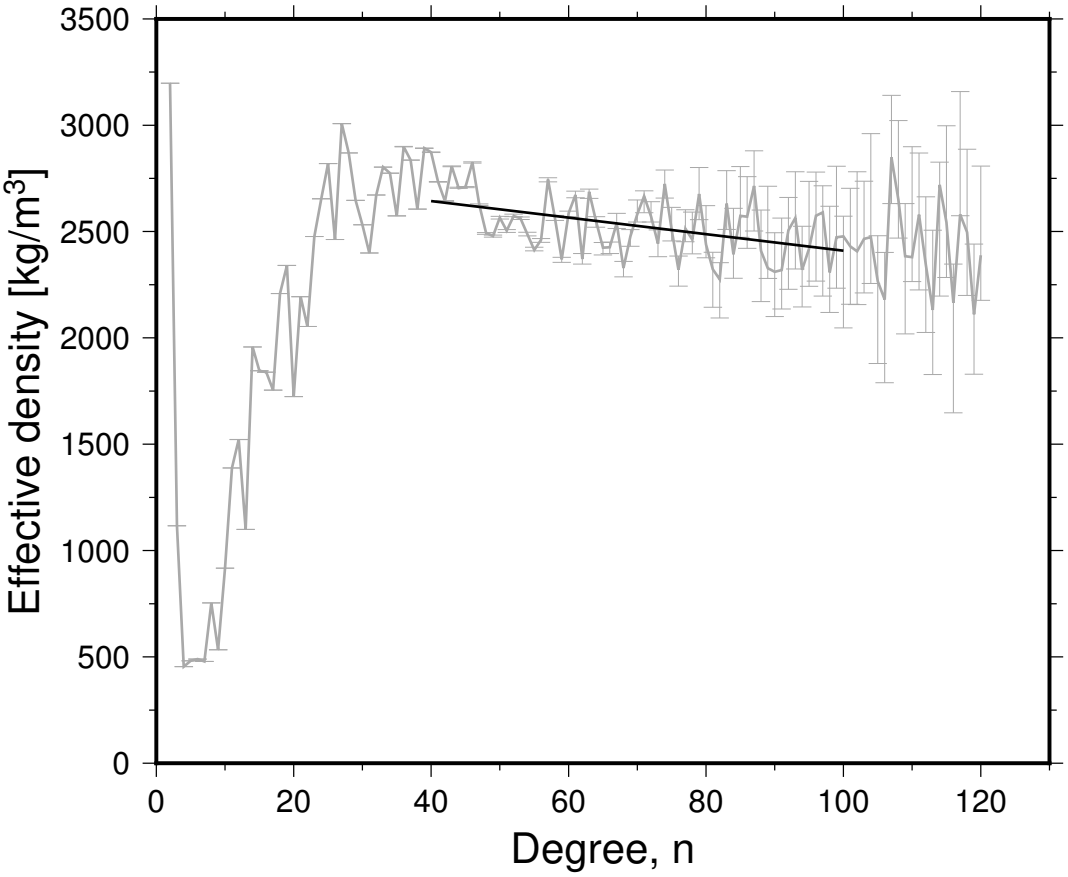

Figure S13: The same as Fig. S12, with the errors now given by the minimum and maximum effective density values per degree from the set of clone models.

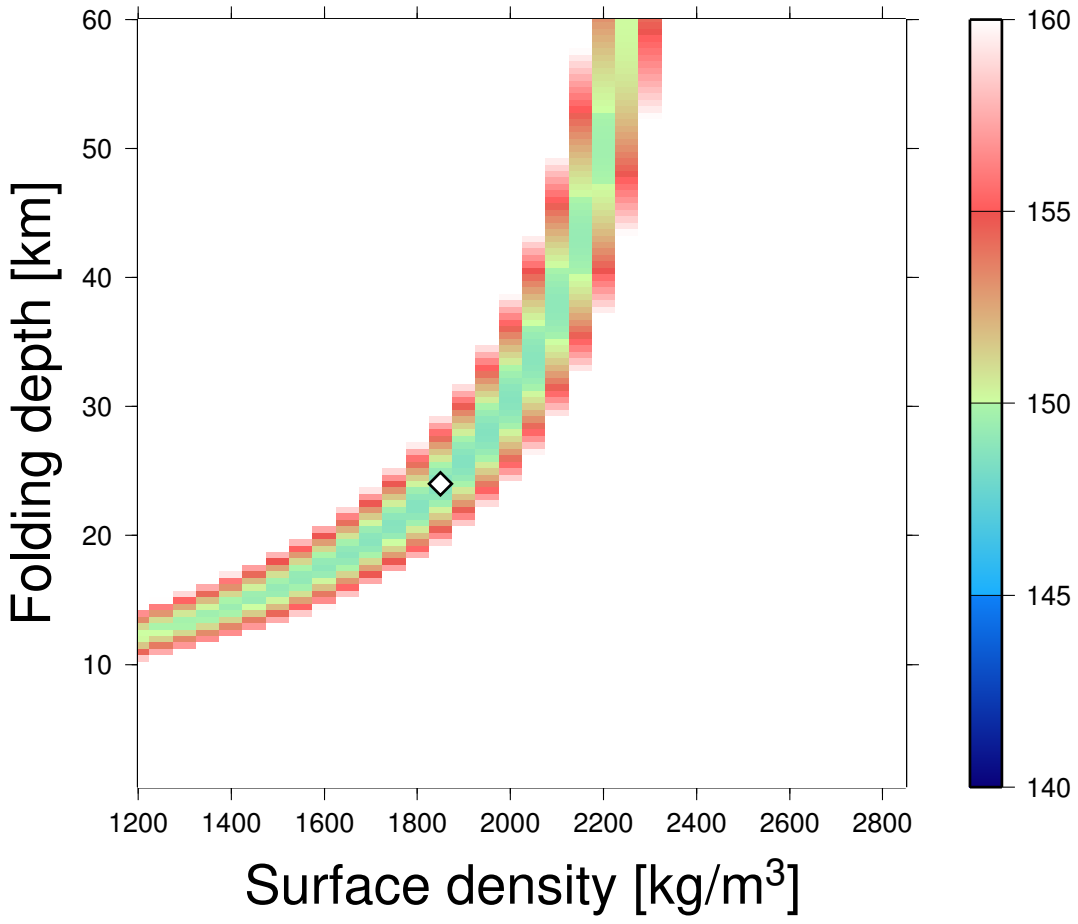

Figure S14: A map of the misfit between analytic models of admittance and the Mars effective density spectrum from our RM1 model. The analytic models assume an exponential increase of density with depth, which can readily be related to pore closure through compaction (*Han et al.*, 2014). Surface density and the value of the e-folding depth were the free parameters. The best-fit model (in a root-mean-square sense, computed between degrees $n = 60$ and $n = 120$) is indicated with a diamond, but the misfit plot shows that a wide variety of models can fit the data equally well.

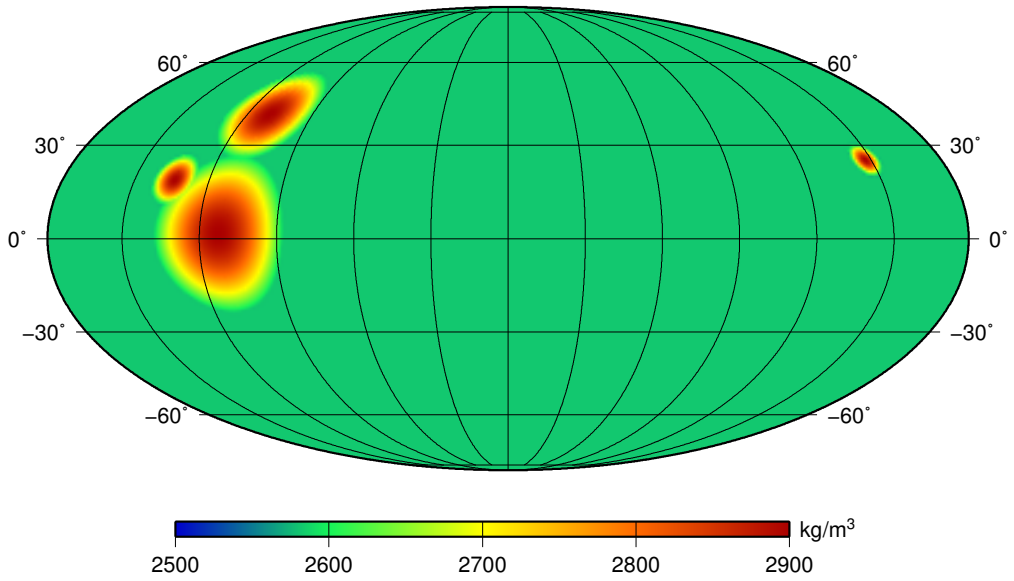

Figure S15: A map of laterally varying crustal density for Mars. The map is in Mollweide projection centered on the prime meridian. Our average value of 2582 kg m^{-3} is used everywhere except for the following volcanic complexes: Alba, Tharsis, Olympus Mons and Elysium Mons. For those, a value of 2900 kg m^{-3} is used (instead of a value close to 3200 kg m^{-3}, for comparison with the standard value, and to avoid convergence issues in the crustal thickness determination process), and the density is transitioned smoothly using a cosine taper and assigned radius around the center (15°, 25°, 8°, and 5° for the aforementioned volcanoes, respectively).

Supplementary References

Besserer, J., F. Nimmo, M. A. Wieczorek, R. C. Weber, W. S. Kiefer, P. J. McGovern, J. C. Andrews-Hanna, D. E. Smith, and M. T. Zuber (2014), GRAIL gravity constraints on the vertical and lateral density structure of the lunar crust, *Geophysical Research Letters*, *41*, 5771–5777, doi:10.1002/2014GL060240.

Genova, A., S. Goossens, F. Lemoine, E. Mazarico, G. Neumann, D. Smith, and M. Zuber (2016), Seasonal and static gravity field of Mars from MGS, Mars Odyssey and MRO radio science, *Icarus*, *272*, 228–245, doi:10.1016/j.icarus.2016.02.050.

Grott, M., and M. A. Wieczorek (2012), Density and lithospheric structure at Tyrrhena Patera, Mars, from gravity and topography data, *Icarus*, *221*, 43–52, doi:10.1016/j.icarus.2012.07.008.

Han, S.-C., N. Schmerr, G. Neumann, and S. Holmes (2014), Global characteristics of porosity and density stratification within the lunar crust from GRAIL gravity and LOLA topography data, *Geophysical Research Letters*, *41*, 1882–1889, doi:10.1002/2014GL059378.

Kaula, W. M. (1966), *Theory of Satellite Geodesy, applications of satellites to Geodesy*, Blaisdell Publishing Company, Waltham, MA., USA.

Konopliv, A. S., W. B. Banerdt, and W. L. Sjogren (1999), Venus Gravity: 180th Degree and Order Model, *Icarus*, *139*, 3–18, doi:10.1006/icar.1999.6086.

Konopliv, A. S., J. K. Miller, W. M. Owen, D. K.Yeomans, J. D. Giorgini, R. Garmier, and J. P. Barriot (2002), A Global Solution for the Gravity Field, Rotation, Landmarks, and Ephemerides of Eros, *Icarus*, *160*, 289–299, doi:10.1006/icar.2002.6975.

Konopliv, A. S., C. F. Yoder, E. M. Standish, D. N. Yuan, and W. L. Sjogren (2006), A global solution for the Mars static and seasonal gravity, Mars orientation, Phobos and Deimos masses, and Mars ephemeris, *Icarus*, *182*, 23–50, doi:10.1016/j.icarus.2005.12.025.

Konopliv, A. S., S. W. Asmar, R. S. Park, B. G. Bills, F. Centinello, A. B. Chamberlin, A. Ermakov, R. W. Gaskell, N. Rambaux, C. A. Raymond, C. T. Russell, D. E. Smith, P. Tricarico, and M. T. Zuber (2014), The Vesta gravity field, spin pole and rotation period, landmark positions, and ephemeris from the Dawn tracking and optical data, *Icarus*, *240*, 103 – 117, doi:10.1016/j.icarus.2013.09.005.

Konopliv, A. S., R. S. Park, and W. M. Folkner (2016), An improved JPL Mars gravity field and orientation from Mars orbiter and lander tracking data, *Icarus*, *274*, 253–260, doi:10.1016/j.icarus.2016.02.052.

Lemoine, F. G., S. Goossens, T. J. Sabaka, J. B. Nicholas, E. Mazarico, D. D. Rowlands, B. D. Loomis, D. S. Chinn, D. S. Caprette, G. A. Neumann, D. E. Smith, and M. T. Zuber (2013), High-degree gravity models from GRAIL primary mission data, *Journal of Geophysical Research*, *118*(8), 1676–1698, doi:10.1002/jgre.20118.

Lemoine, F. G., S. Goossens, T. J. Sabaka, J. B. Nicholas, E. Mazarico, D. D. Rowlands, B. D.

Loomis, D. S. Chinn, G. A. Neumann, D. E. Smith, and M. T. Zuber (2014), GRGM900C: A degree-900 lunar gravity model from GRAIL primary and extended mission data, *Geophysical Research Letters*, *41*(10), 3382–3389, doi:10.1002/2014GL060027.

Mazarico, E., A. Genova, S. Goossens, F. G. Lemoine, G. A. Neumann, M. T. Zuber, D. E. Smith, and S. C. Solomon (2014), The Gravity Field, Orientation, and Ephemeris of Mercury from MESSENGER Observations after Three Years in Orbit, *Journal of Geophysical Research Planets*, *119*(12), 2417–2436, doi:10.1002/2014JE004675.

Neumann, G. A., M. T. Zuber, M. A. Wieczorek, P. J. McGovern, F. G. Lemoine, and D. E. Smith (2004), Crustal structure of Mars from gravity and topography, *Journal of Geophysical Research*, *109*(E08002), doi:10.1029/2004JE002262.

Ojha, L., K. Lewis, and S. Karunatillake (2017), Volcanic Origin of Medusae Fossae Formation from Gravity and Topography Data, in *48th Lunar and Planetary Science Conference, abstract 2475*, The Woodlands, Texas, USA.

Seber, G. A. F., and C. J. Wild (1989), *Nonlinear Regression*, Wiley-Interscience.

Wieczorek, M. A. (2008), Constraints on the composition of the martian south polar cap from gravity and topography, *Icarus*, *196*, 506–517, doi:10.1016/j.icarus.2007.10.026.

Wieczorek, M. A. (2015), Gravity and Topography of the Terrestrial Planets, in *Treatise on Geophysics (Second Edition)*, edited by G. Schubert, pp. 153–193, Elsevier, Oxford, doi:10.1016/B978-0-444-53802-4.00169-X.

Wieczorek, M. A., and M. Meschede (2015), SHTOOLS - Tools for working with spherical harmonics (v3.0), ZENODO, doi:10.5281/zenodo.15967.

Wieczorek, M. A., and R. J. Phillips (1998), Potential anomalies on a sphere: Applications to the thickness of the lunar crust, *Journal of Geophysical Research*, *103*(E1), 1715–1724, doi:10.1029/97JE03136.

Wieczorek, M. A., and F. J. Simons (2005), Localized spectral analysis on the sphere, *Geophysical Journal International*, *162*, 655–675, doi:10.1111/j.1365-246X.2005.02687.x.

Wieczorek, M. A., G. A. Neumann, F. Nimmo, W. S. Kiefer, G. J. Taylor, H. J. Melosh, R. J. Phillips, S. C. Solomon, J. C. Andrews-Hanna, S. W. Asmar, A. S. Konopliv, F. G. Lemoine, D. E. Smith, M. M. Watkins, J. G. Williams, and M. T. Zuber (2013), The Crust of the Moon as Seen by GRAIL, *Science*, *339*(6120), 671–675, doi:10.1126/science.1231530.