A&A 620, A59 (2018)
https://doi.org/10.1051/0004-6361/201832987
© ESO 2018

Astronomy & Astrophysics

Solving linear equations with messenger-field and conjugate gradient techniques: An application to CMB data analysis

J. Papež[1], L. Grigori[1], and R. Stompor[2]

[1] INRIA Paris, Sorbonne Université, Univ. Paris-Diderot SPC, CNRS, Laboratoire Jacques-Louis Lions, équipe ALPINES, France
e-mail: jan.papez@inria.fr
[2] AstroParticule et Cosmologie, Univ. Paris Diderot, CNRS/IN2P3, CEA/Irfu, Obs. de Paris, Sorbonne Paris Cité, France

Received 9 March 2018 / Accepted 27 September 2018

ABSTRACT

We discuss linear system solvers invoking a messenger-field and compare them with (preconditioned) conjugate gradient approaches. We show that the messenger-field techniques correspond to fixed point iterations of an appropriately preconditioned initial system of linear equations. We then argue that a conjugate gradient solver applied to the same preconditioned system, or equivalently a preconditioned conjugate gradient solver using the same preconditioner and applied to the original system, will in general ensure at least a comparable and typically better performance in terms of the number of iterations to convergence and time-to-solution. We illustrate our conclusions with two common examples drawn from the cosmic microwave background (CMB) data analysis: Wiener filtering and map-making. In addition, and contrary to the standard lore in the CMB field, we show that the performance of the preconditioned conjugate gradient solver can depend significantly on the starting vector. This observation seems of particular importance in the cases of map-making of high signal-to-noise ratio sky maps and therefore should be of relevance for the next generation of CMB experiments.

Key words. methods: numerical – cosmic background radiation

1. Introduction

Studies of the cosmic microwave background (CMB) anisotropies have been driving progress in our understanding of the universe for nearly a quarter of a century. The current forefront of the CMB research is the characterization of polarization properties of the CMB anisotropies. The next generation of the CMB observatories has been, and is, designed to ensure that the scientific potential of this new probe is fully exploited. This calls for advanced, high-performance data analysis techniques applicable to enormous data sets which will be collected by these new observatories.

The analysis of data from CMB observations commonly involves solutions of large, structured linear systems of equations. Two typical and important examples of such systems are map-making and Wiener-filter systems of equations (see, e.g., Janssen & Gulkis 1992; Bunn et al. 1994, respectively, for early pioneering work and Poletti et al. 2017; Seljebotn et al. 2017 for examples of more recent applications). These systems are solved either as a stand-alone task or as part of a more involved process, such as a power spectrum estimation, which commonly requires multiple solutions of such systems. In this work we study, from a theoretical and practical perspective, two specific algorithms for solving such systems of equations: a preconditioned conjugate gradient (PCG) approach and a messenger-field (MF) technique. Both these approaches have been applied in the context of the applications considered here. Of the two, the PCG approach has been more popular and more broadly used to date. Nevertheless, it has been argued in a number of recent papers (e.g., Elsner & Wandelt 2013; Ramanah et al. 2017; Huffenberger & Næss 2018) that the messenger-field approach can be highly efficient for both these applications and can deliver performance in some cases exceeding that of some specific PCG approaches, while at the same time being more generally feasible and straightforward to implement and apply (Ramanah et al. 2017; Huffenberger & Næss 2018; Huffenberger 2018). We note that it is the combination of all these features that makes the MF approach potentially attractive. Indeed, performance of the PCG technique is hinged on a choice of a preconditioner matrix, $\mathbf{M}$, and while very efficient preconditioners can be constructed, in principle outperforming other methods, typically the construction quickly becomes difficult and potentially prohibitive from a computational point of view.

Specifically, let us consider a linear system of equations,

$$\mathbf{A}\,\mathbf{x} = \mathbf{b}, \tag{1}$$

where the system matrix $\mathbf{A}$ is symmetric and positive definite (SPD). Instead of solving this equation directly, in the PCG approach one solves the preconditioned system,

$$\mathbf{M}^{-1}\,\mathbf{A}\,\mathbf{x} = \mathbf{M}^{-1}\,\mathbf{b}, \tag{2}$$

applying the conjugate gradient (CG) technique (Golub & Van Loan 1996). If the preconditioner is chosen in a way such that $\mathbf{M}^{-1}\,\mathbf{A}$ is better conditioned than the original system matrix $\mathbf{A}$, then the solution can be derived in (often significantly) fewer iterations. Hereafter, we define the condition number κ as

$$\kappa(\mathbf{A}) \equiv \|\mathbf{A}\|_2 \cdot \|\mathbf{A}^{-1}\|_2, \tag{3}$$

where $||\mathbf{A}||_2$ is the spectral norm[1] of the matrix $\mathbf{A}$. For a good preconditioner $\kappa(\mathbf{A}) \gg \kappa(\mathbf{M}^{-1}\,\mathbf{A}) \geq 1$. A preconditioner is therefore better when its inverse succeeds in capturing more essential features of the inverse system matrix $\mathbf{A}^{-1}$, which cannot be computed directly by assumption. The choice of a preconditioner is a key factor in determining the performance of a PCG solver. There exist both generic and case-specific approaches proposed for their construction. Moreover, for many advanced preconditioners, significant savings in terms of the number of iterations to a solution come at the cost of an overhead related to their construction and/or application on every step of iteration.

While these observations make the method comparison cumbersome and potentially limited to very specific cases and concrete implementations, the question of which class of methods is more promising in ensuring sufficient performance for forthcoming data sets, for example in the context of the CMB field, is of actual practical importance. This is the question we tackle in this work in the context of the MF and PCG solvers.

Our methodology is as follows. We first show that any MF method applied to a linear system involves preconditioning of the original set of equations with a specific preconditioner. Then we argue on theoretical grounds, and later demonstrate using a number of study cases that the corresponding PCG algorithm is at least as efficient as, and often much better than, the MF technique, while featuring similar computational complexity. Combining the MF method with a so-called cooling technique can further improve its performance at least in some cases. Nonetheless, this does not seem to affect the overall assessment at least for the specific cooling prescriptions studied in this paper and motivated by earlier work. Consequently, while the MF technique may still provide an interesting alternative in some specific applications, in general the PCG approach seems more promising and can be a better first choice. Though we demonstrate our conclusions using specific examples from applications to the Wiener filter and map-making procedures, we expect them to hold more generally.

This paper is organized as follows. In Sect. 2 we provide a general discussion of the messenger-field technique as a more general class of solvers and compare it with the PCG solvers as far as its convergence and computational aspects are concerned. Later, we illustrate the general conclusion of this section with the help of numerical experiments applied to simulated CMB data, involving applications of both these techniques to polarized Wiener filtering (Sect. 3) and map-making (Sect. 4). We conclude in Sect. 5. Some more technical considerations are deferred to the Appendices.

2. Messenger-field iterative solver

In this section we first present a consistent, general algebraic framework, which encompasses, as specific cases, all implementations of the basic messenger field solver proposed to date in the literature. Subsequently, in Sect. 2.2 we describe the cooling technique proposed to improve the performance of the messenger method, and in Sects. 2.3 and 2.4 we discuss the general properties of this broad class of solvers, contrasting them with those of the PCG technique. We develop this general discussion, referring to the specific solvers developed in the context of the applications described in the following sections of this paper.

2.1. Basic approach

Let us consider a system of linear equations as in Eq. (1). In general, the messenger-field approach involves a split of the system matrix $\mathbf{A}$, such as

$$\mathbf{A} \equiv \mathbf{C} - \mathbf{D} = \mathbf{C}\,(\mathbf{I} - \mathbf{C}^{-1}\,\mathbf{D}), \tag{4}$$

where $\mathbf{C}$ is invertible by construction and its inverse is easy to compute. $\mathbf{I}$ is an identity matrix. After multiplying Eq. (1) by $\mathbf{C}^{-1}$ from the left (which corresponds to preconditioning the original system), we get the system

$$(\mathbf{I} - \mathbf{C}^{-1}\,\mathbf{D})\,\mathbf{x} = \mathbf{C}^{-1}\,\mathbf{b}. \tag{5}$$

The MF method introduces an extra data object $\mathbf{t}$, the messenger field, which can be defined as

$$\mathbf{t} \equiv \mathbf{D}\,\mathbf{x} + \mathbf{b}, \tag{6}$$

meaning that Eq. (5) can be represented as

$$\mathbf{t} = \mathbf{D}\,\mathbf{x} + \mathbf{b}, \tag{7}$$
$$\mathbf{x} = \mathbf{C}^{-1}\,\mathbf{t}. \tag{8}$$

This can be used to define an iterative scheme

$$\mathbf{t}^{(i+1)} = \mathbf{D}\,\mathbf{x}^{(i)} + \mathbf{b}, \tag{9}$$
$$\mathbf{x}^{(i+1)} = \mathbf{C}^{-1}\,\mathbf{t}^{(i+1)}. \tag{10}$$

We note that the messenger field $\mathbf{t}$ introduced in this way is a dummy object. Therefore, barring some implementational advantages, the equations above are equivalent to a reduced system from which the messenger field has been explicitly eliminated and which can be directly derived from Eq. (5). This reads,

$$\mathbf{x} = \mathbf{C}^{-1}\,\mathbf{D}\,\mathbf{x} + \mathbf{C}^{-1}\,\mathbf{b}, \tag{11}$$

and the corresponding iterative scheme, see also (Elsner & Wandelt 2013; Huffenberger & Næss 2018), is given by

$$\mathbf{x}^{(i+1)} = \mathbf{C}^{-1}\,\mathbf{D}\,\mathbf{x}^{(i)} + \mathbf{C}^{-1}\,\mathbf{b}. \tag{12}$$

This is a fixed-point iteration scheme (e.g., Saad 2003) and its derivation is analogous to the derivation of the classical iteration methods that also rely on the splitting of the system matrix as in Eq. (4). The Jacobi iterative method takes $\mathbf{C}$ as the diagonal of $\mathbf{A}$, while in the Gauss–Seidel method $\mathbf{C}$ is equal to the lower triangular part (including the diagonal) of $\mathbf{A}$.

We emphasize that whether we choose to implement the single equation version as in Eq. (12) or the double equation one as in Eqs. (9) and (10), the result will be the same to within numerical precision as in both these cases we solve the same linear system, Eq. (5), performing equivalent iterations. Consequently, the messenger-field approach is a fixed-point iteration technique applied to a preconditioned system in Eq. (5). However, this equation can be solved using other means, such as for instance a conjugate gradient (CG) approach, which is typically more efficient than the fixed-point iterations (see, e.g., Sects. 5.5 and 2.3 of Liesen & Strakoš (2013)) Moreover, solving Eq. (5) with the help of the CG technique is equivalent to solving the initial set of equations, Eq. (1), using a PCG technique with the preconditioner set to $\mathbf{M} \equiv \mathbf{C}$. In cases when the fixed-point method is expected to converge very efficiently, that is, when

[1] The spectral norm of $\mathbf{A}$ is equal to the largest singular value of $\mathbf{A}$ defined as the square root of the largest eigenvalue of $\mathbf{A}^\dagger\mathbf{A}$, where $\dagger$ denotes a Hermitian conjugate. If matrix $\mathbf{A}$ is *normal*, i.e., $\mathbf{A}\,\mathbf{A}^\dagger = \mathbf{A}^\dagger\,\mathbf{A}$, then $\mathbf{A}$ can be diagonalized with the help of a similarity operation employing a unitary matrix and the spectral norm of $\mathbf{A}$ is equal to its largest (in magnitude) eigenvalue and the condition number to the ratio of the largest and the smallest eigenvalues.

$\mathbf{A} \simeq \mathbf{C}$, the PCG solver will also perform well since $\mathbf{C}^{-1}\,\mathbf{A} \simeq \mathbf{I}$, a hallmark of a good preconditioner. Similarly, the MF solver based on the split involving a good preconditioner will likely be efficient. From a computational point of view, both techniques require multiple applications of the inverse preconditioner $\mathbf{M}^{-1}$ to a vector, thus resulting in similar numerical cost.

The main message of this section is that the messenger-field method involves fixed-point iterations applied to a preconditioned system of linear equations. Its performance is determined by an adopted split of the system matrix, which also defines the preconditioner applied to precondition the initial system. This preconditioner can be used alternately in a PCG solver employed to directly solve the initial system, and is expected to ensure performance as good as or better than that of the MF technique, as far as the number of iterations as well as time to convergence are concerned.

2.2. Cooling technique

The convergence of the fixed-point method, Eq. (14), depends on the components of the initial error $\mathbf{x} - \mathbf{x}^{(0)}$ in the invariant subspaces associated with the eigenvalues of $\mathbf{C}^{-1}\,\mathbf{D}$, especially with the dominant (largest) ones. The cooling technique proposed in Elsner & Wandelt (2013) aims at providing, iteratively, a good initial guess $\mathbf{x}^{(0)}$. In the general setting considered above, the cooling technique replaces the original problem, Eq. (1), represented in a split form as in Eq. (4) by

$$\mathbf{A}(\lambda)\mathbf{x}(\lambda) = \mathbf{b}(\lambda), \qquad \mathbf{A}(\lambda) = \mathbf{C}(\lambda) - \mathbf{D}(\lambda), \tag{13}$$

where the cooling parameter λ is defined so that: (a) for $\lambda = 1$, the above problem is equivalent to the original problem in Eqs. (1) and (4); and (b) for $\lambda \to \infty$, $\mathbf{D}(\lambda) \to 0$ and $(\mathbf{C}(\lambda))^{-1}\mathbf{D}(\lambda) \to 0$. In the cooling method, λ is progressively adapted in the course of the iterations with its value gradually decreasing from an initial and typically rather large value down to 1. While no general prescription is given in the literature, it has been claimed (Elsner & Wandelt 2013; Huffenberger & Næss 2018) that at least in some applications significant gains can be derived as compared to the fixed-point iterations, if the rate of change of λ is appropriately tuned. In general, an MF method combined with the cooling is no longer a fixed-point method. However, as this is often the case, if λ does not change with each iteration but rather is kept constant for some number of iterations before assuming a new value, the iterations for each of the fixed values of the parameters are fixed-point (though not of the original system). In such cases, for large values of λ it is expected that an accurate solution of the modified system, Eq. (13), that is, $\mathbf{x}(\lambda)$, can be recovered within a few (fixed-point) iterations. Naturally, $\mathbf{x}(\lambda)$ can be far from the desired solution of the actual system, $\mathbf{x}(1)$, however, it can be a good starting vector for the next round of fixed-point iterations, this time with a smaller value of λ. The relative performance of the cooling method compared to that of the PCG solver of the initial equation, i.e., with $\lambda = 1$, is unclear, and the freedom in defining the rate at which λ is changed makes the mathematical analysis of this method difficult; its potential advantages over others are therefore also difficult to anticipate. Consequently, in this work we resort to numerical experiments to investigate the pros and cons of this technique in the specific cases of interest (Sects. 3 and 4).

We note however that a PCG solver could be used instead of the fixed-point iterations within the cooling scheme. Though, the fixed-point iterations would still be preferable whenever the value of λ is adjusted after each iteration, or every few iterations, for example, as in the cooling scheme proposed in Sect. 2.2 of Elsner & Wandelt (2013). However, in the cases when the value of λ is kept unchanged over a number of iterations, as in the numerical experiments presented in Huffenberger & Næss (2018) and in Ramanah et al. (2017), replacing the fixed-point iterations by a PCG method is expected to result in some performance gain accumulated from all the gains obtained from the solutions for a fixed value of λ.

For clarity, hereafter we use the term "messenger-field method" to denote a method which implements the basic MF algorithm as defined in Sect. 2.1. Whenever cooling is involved, be it combined with the MF method or the PCG one, we explicitly point this out; for example, we refer to the "cooled MF" or the "PCG with cooling", and vice versa.

2.3. Convergence

The convergence properties of the classic, fixed-point iteration methods have been studied extensively in the literature (see, e.g., Sects. 4.1–4.2 of Saad 2003, or Sect. 10.1 of Golub & Van Loan 1996). Given our discussion in Sect. 2.1 those results can be directly applied to the messenger field technique. In particular, from Eqs. (11) and (12) (see also, Elsner & Wandelt 2013) the error of the ith approximation satisfies the following relation.

$$\begin{aligned} \epsilon^{(i)} \;&\equiv\; \mathbf{x} - \mathbf{x}^{(i)} \;=\; \mathbf{C}^{-1}\,\mathbf{D}\,(\mathbf{x} - \mathbf{x}^{(i-1)}) \\ &\equiv\; \mathbf{C}^{-1}\,\mathbf{D}\,\epsilon^{(i-1)} \;=\; [\mathbf{C}^{-1}\,\mathbf{D}]^{i}\,\epsilon^{(0)}. \end{aligned} \tag{14}$$

This implies (see, e.g., Golub & Van Loan 1996; Saad 2003) that $\|\epsilon^{(i)}\|$ converges asymptotically to zero as long as the spectral radius of $\mathbf{C}^{-1}\,\mathbf{D}$ is smaller than unity. Here $\|\cdot\|$ denotes the Euclidean norm, and the spectral radius, hereafter denoted by $\rho(\cdot)$, is defined as the largest (in magnitude) eigenvalue of the matrix. This observation generalizes to other norms given their equivalence on finite dimensional spaces; see Appendix B for more details.

If matrix $\mathbf{C}^{-1}\,\mathbf{D}$ is also normal, then from (14) it follows that,

$$\|\epsilon^{(i)}\| \;\leq\; \rho(\mathbf{C}^{-1}\,\mathbf{D})\,\|\epsilon^{(i-1)}\|, \tag{15}$$

in the Euclidean norm. Therefore, in this case the convergence is not only asymptotic but also monotonic. The normality of $\mathbf{C}^{-1}\,\mathbf{D}$ is also typically necessary (see, e.g., Sect. 4.1.6 of Björck 2015 for relevant examples). Consequently, in general some care may need to be exercised in choosing a specific split of the system matrix, Eq. (4), to ensure that it satisfies both these conditions. This is indeed the case for the Wiener filter application (see, Elsner & Wandelt 2013 and Appendix B), as the spectral radius of $\mathbf{C}^{-1}\,\mathbf{D}$ is always smaller than 1, assuming that the corresponding system matrix $\mathbf{A}$ is non-singular (see below) and $\mathbf{C}^{-1}\,\mathbf{D}$ is normal. For the map-making application in the rendition of Huffenberger & Næss (2018), arguments similar to those given in Appendix B can be used to show that also in this case $\rho(\mathbf{C}^{-1}\,\mathbf{D}) < 1$ for non-singular systems, however the normality of $\mathbf{C}^{-1}\,\mathbf{D}$ remains unclear at this stage.

The character of the convergence will in general depend on the choice of the norm. We show here that it remains monotonic in the $\mathbf{A}$-norm (often called *energy norm)* of the error, if $\mathbf{C}^{-1}$ and $\mathbf{D}$ are real and symmetric as is indeed the case in the applications studied here. The $\mathbf{A}$-norm is hereafter defined as,

$$\|\mathbf{x} - \mathbf{x}^{(i)}\|_{\mathbf{A}} \equiv \left((\mathbf{x} - \mathbf{x}^{(i)})^{T}\mathbf{A}(\mathbf{x} - \mathbf{x}^{(i)})\right)^{1/2}. \tag{16}$$

This is one of the norms we use in the follow-up numerical examples.

Using Eq. (14) and $\|\mathbf{v}\|_{\mathbf{A}} = \|\mathbf{A}^{1/2}\,\mathbf{v}\|$, we obtain,

$$\|\mathbf{x} - \mathbf{x}^{(i)}\|_{\mathbf{A}} \;\le\; \|\mathbf{B}\|_2 \cdot \|\mathbf{x} - \mathbf{x}^{(i-1)}\|_{\mathbf{A}}, \tag{17}$$

where,

$$\mathbf{B} \;\equiv\; \mathbf{A}^{1/2}\,(\mathbf{C}^{-1}\mathbf{D})\,\mathbf{A}^{-1/2}, \tag{18}$$

and $\|\cdot\|_2$ denotes the spectral norm, defined as in Eq. (3).

To ensure monotonic convergence of the iterative scheme, Eq. (12), in the energy norm it is therefore enough to require that

$$\|\mathbf{B}\|_2 < 1. \tag{19}$$

If $\rho(\mathbf{C}^{-1}\,\mathbf{D}) < 1$, then Eq. (19) is satisfied whenever matrix $\mathbf{B}$ is normal, that is, $\mathbf{B}\,\mathbf{B}^{\dagger} = \mathbf{B}^{\dagger}\,\mathbf{B}$, and therefore it holds that

$$\|\mathbf{B}\|_2 = \rho\,(\mathbf{B}) \;=\; \rho(\,\mathbf{A}^{1/2}\,(\mathbf{C}^{-1}\mathbf{D})\,\mathbf{A}^{-1/2}) \;=\; \rho\left(\mathbf{C}^{-1}\mathbf{D}\right). \tag{20}$$

Here the leftmost equality uses the fact that a normal matrix can be diagonalized using a unitary matrix, and the rightmost follows from the fact that a non-singular similarity transformation, here with $\mathbf{A}^{1/2}$, preserves the eigenvalues.

Assuming that $\mathbf{A}$ is real and symmetric and observing from Eq. (4) that $\mathbf{A}\,(\mathbf{C}^{-1}\,\mathbf{D}) = (\mathbf{D}\,\mathbf{C}^{-1})\,\mathbf{A}$, we can write,

$$\begin{aligned}
\mathbf{B}^{\dagger}\,\mathbf{B} &= \mathbf{A}^{-1/2}\,(\mathbf{C}^{-1}\mathbf{D})^{\dagger}\,\mathbf{A}\,\mathbf{C}^{-1}\,\mathbf{D}\,\mathbf{A}^{-1/2} \\
&= \mathbf{A}^{-1/2}\,(\mathbf{C}^{-1}\mathbf{D})^{\dagger}\,\mathbf{D}\,\mathbf{C}^{-1}\,\mathbf{A}^{1/2}, &(21)\\
\mathbf{B}\,\mathbf{B}^{\dagger} &= \mathbf{A}^{1/2}\,(\mathbf{C}^{-1}\mathbf{D})\,\mathbf{A}^{-1}\,(\mathbf{C}^{-1}\mathbf{D})^{\dagger}\,\mathbf{A}^{1/2} \\
&= \mathbf{A}^{-1/2}\,\mathbf{D}\,\mathbf{C}^{-1}\,(\mathbf{C}^{-1}\mathbf{D})^{\dagger}\,\mathbf{A}^{1/2}, &(22)
\end{aligned}$$

and therefore, $\mathbf{B}$ is normal if and only if $(\mathbf{C}^{-1}\mathbf{D})^{\dagger}\,\mathbf{D}\,\mathbf{C}^{-1} = \mathbf{D}\,\mathbf{C}^{-1}\,(\mathbf{C}^{-1}\mathbf{D})^{\dagger}$. This is satisfied, for example, whenever $\mathbf{C}$ and $\mathbf{D}$ are real and symmetric as indeed is the case in the setting of Elsner & Wandelt (2013) and Huffenberger & Næss (2018).

Consequently, in both these applications we expect monotonic convergence of the MF errors in the energy norm. This is analogous to the PCG technique where the error is bound to decrease monotonically in the energy norm. This may not however be the case for other norms or the residuals.

The convergence rate of the MF solver is then determined by the eigenspectrum of $\mathbf{C}^{-1}\mathbf{D}$. In particular the eigenmode with the largest eigenvalue, that is, the one closest to 1, will be the slowest to converge.

If the system matrix, $\mathbf{A}$, is singular then

$$\rho(\mathbf{C}^{-1}\,\mathbf{D}) \;=\; 1. \tag{23}$$

This is because if $\mathbf{x}$ denotes a singular eigenvector of $\mathbf{A}$, that is, $\mathbf{A}\,\mathbf{x} = 0$, and $\mathbf{x} \neq 0$ then,

$$\mathbf{A}\,\mathbf{x} \;=\; \mathbf{C}\,(\mathbf{x} - \mathbf{C}^{-1}\,\mathbf{D}\,\mathbf{x}) = 0, \tag{24}$$

and hence

$$\mathbf{C}^{-1}\,\mathbf{D}\,\mathbf{x} \;=\; \mathbf{x}, \tag{25}$$

and $\mathbf{x}$ is also an eigenvector of $\mathbf{C}^{-1}\,\mathbf{D}$ but with a unit eigenvalue. In such cases, the convergence of the MF solver will typically stall with the norm of the residuals saturating on a level depending on the right-hand side of the system as well as the initial guess. This behavior is analogous to that of other solvers, such as PCG, and it simply reflects the fact that if $\mathbf{A}$ is singular, then there is no unique solution to the linear system.

We assume from now on that the problem is non-singular and show that the PCG method is typically superior to, and never worse than, the fixed-point method in terms of minimizing the energy norm of the error. We first recall key properties of the (P)CG approach; see, for example, Saad (2003), Lemma 6.28.

Let $\mathbf{x}^{(\mathrm{CG},i)}$ be the ith approximation given by the CG method for solving $\mathbf{A}\mathbf{x} = \mathbf{b}$ with the initial guess $\mathbf{x}^{(0)}$. Subsequently,

$$\mathbf{x} - \mathbf{x}^{(\mathrm{CG},i)} = \hat{\varphi}_i(\mathbf{A})(\mathbf{x} - \mathbf{x}^{(0)}), \tag{26}$$

where $\hat{\varphi}_i$ is a polynomial with $\deg(\hat{\varphi}_i) \le i$, $\hat{\varphi}_i(0) = 1$, which we write succinctly as $\hat{\varphi}_i \in \mathcal{P}_0^i$, and,

$$\|\hat{\varphi}_i(\mathbf{A})(\mathbf{x} - \mathbf{x}^{(0)})\|_{\mathbf{A}} = \min_{\hat{\psi}_i \in \mathcal{P}_0^i} \|\hat{\psi}_i(\mathbf{A})(\mathbf{x} - \mathbf{x}^{(0)})\|_{\mathbf{A}}. \tag{27}$$

Similarly, when $\mathbf{x}^{(\mathrm{PCG},i)}$ is the ith approximation given by the PCG method for solving the system $\mathbf{A}\mathbf{x} = \mathbf{b}$ preconditioned by $\mathbf{C}$, using the initial guess $\mathbf{x}^{(0)}$, we have

$$\mathbf{x} - \mathbf{x}^{(\mathrm{PCG},i)} = \varphi_i(\mathbf{C}^{-1}\mathbf{A})(\mathbf{x} - \mathbf{x}^{(0)}), \tag{28}$$

with $\varphi_i \in \mathcal{P}_0^i$ and

$$\|\varphi_i(\mathbf{C}^{-1}\mathbf{A})(\mathbf{x} - \mathbf{x}^{(0)})\|_{\mathbf{A}} = \min_{\psi_i \in \mathcal{P}_0^i} \|\psi_i(\mathbf{C}^{-1}\mathbf{A})(\mathbf{x} - \mathbf{x}^{(0)})\|_{\mathbf{A}}. \tag{29}$$

Let us now consider the fixed-point method as defined in Eq. (12) assuming the same initial guess, $\mathbf{x}^{(0)}$. From Eqs. (29) and (14),

$$\begin{aligned}
\|\mathbf{x} - \mathbf{x}^{(\mathrm{PCG},i)}\|_{\mathbf{A}} &= \|\varphi_i(\mathbf{C}^{-1}\mathbf{A})(\mathbf{x} - \mathbf{x}^{(0)})\|_{\mathbf{A}} &(30)\\
&\le \|(\mathbf{C}^{-1}\mathbf{D})^i(\mathbf{x} - \mathbf{x}^{(0)})\|_{\mathbf{A}} = \|\mathbf{x} - \mathbf{x}^{(i)}\|_{\mathbf{A}}, &(31)
\end{aligned}$$

as $(\mathbf{C}^{-1}\mathbf{D})^i = (\mathbf{I} - \mathbf{C}^{-1}\mathbf{A})^i$ and $\psi_i(x) \equiv (1-x)^i \in \mathcal{P}_0^i$. This means that, in terms of the energy norm of the error, the PCG method converges at least as fast as the fixed-point method. In practice, one can however expect significantly faster convergence, as suggested by Eq. (29). On the other hand, as emphasized earlier, the performance of the MF solvers can be improved by invoking the cooling technique. The convergence of the cooled MF approach is more difficult to study theoretically. Even in the cases when the cooling parameter, λ, is kept constant over some number of iterations, and the method performs fixed-point iterations within each such interval, these are fixed iterations of the modified, not the original, system and the results concerning the error specified earlier in this section apply only when replacing $\mathbf{x}$ by the modified solution $\mathbf{x}(\lambda)$ (and the energy norm $\|\cdot\|_{\mathbf{A}}$ by $\|\cdot\|_{\mathbf{A}(\lambda)}$). The convergence of the iterates to the true solution $\mathbf{x}$ then should be properly discussed and justified for a particular MF application and/or cooling scheme. In the absence of theoretical results concerning this last method we assess the relative merits of the different solvers via numerical experiments. This is described in the follow-up sections.

2.4. Computational complexity

In actual applications, the computational and memory cost per iteration is often as important as algorithmic efficiency. From this perspective, the fixed-point scheme, Eq. (12), is the cheapest method as it requires an evaluation of $\mathbf{C}^{-1}\mathbf{D}\mathbf{x}^{(i)}$ only once per iteration and storing of only two vectors, $\mathbf{x}^{(i)}$, $\mathbf{C}^{-1}\mathbf{b}$. The PCG method requires more memory needed to store up to four or five vectors, depending on the implementation, and each iteration requires two additional inner products plus some scalar multiplications and vector updates. However, typically, and in particular in the applications considered in this paper, the most time-consuming operations are the multiplications by matrix $\mathbf{A}$ and by $\mathbf{C}^{-1}$, rendering these additional costs mostly irrelevant.

As an example, in Appendix A we describe an implementation of the PCG algorithm in the context of the Wiener filter that allows for a single PCG iteration to be performed, with a computational cost comparable to the cost of one fixed-point iteration, Eq. (12).

We can further capitalize on using the PCG method whenever the relative residual or an error measure corresponding to the **A**-norm of the error need to be frequently evaluated; in the extreme case, at each iteration. The residual $\mathbf{r}^{(i)}$ is updated on each PCG iteration and it is therefore at our disposal; this is not the case for the fixed-point iterations, Eq. (12). Similarly, there is a numerically stable way to evaluate the problem-related error measure corresponding to the **A**-norm of the error; see also Appendix A. This evaluation involves only scalar quantities that are already at our disposal during the PCG iterations.

We conclude that in terms of time per iteration, both approaches, the MF and the corresponding PCG, are comparable, and therefore the number of iterations to convergence is a sufficient comparison metric.

3. Application to Wiener filtering

3.1. The problem

Let us consider a sky map $\mathbf{m}$ composed of a sky signal $\mathbf{s}$ and some noise $\mathbf{n}$ due to our instrument, thus

$$\mathbf{m} = \mathbf{s} + \mathbf{n}. \tag{32}$$

We assume that the sky signal is Gaussian over an ensemble of sky realizations with zero mean and known covariance given by $\mathbf{S}$. The noise is also Gaussian with zero mean and the covariance given by $\mathbf{N}$ over the ensemble of noise realizations. We further assume that the noise is uncorrelated and therefore its covariance $\mathbf{N}$ is block-diagonal. The minimum variance estimate of the sky signal, that is, its Wiener filter, is then given by (e.g., Bunn et al. 1994),

$$\mathbf{s}_{\rm WF} = (\mathbf{S}^{-1} + \mathbf{N}^{-1})^{-1}\,\mathbf{N}^{-1}\,\mathbf{m}. \tag{33}$$

Computing the Wiener filter of the measured map, $\mathbf{m}$, requires an inversion of the system matrix, $\mathbf{S}^{-1} + \mathbf{N}^{-1}$. As modern CMB maps may contain up to many millions of pixels this task can indeed be daunting. This is because in general there is no obvious domain in which both the signal and noise covariances are simultaneously diagonal. Indeed, the signal covariance $\mathbf{S}$ is diagonal in the harmonic domain, where the pixel-domain map $\mathbf{m}$ is described by a vector of coefficients $\mathbf{m}_{\ell m}$ obtained as a result of a spherical harmonic transform applied to the map, while the noise covariance is diagonal in the pixel domain, and only diagonal in the harmonic one if the noise is homogeneous, which is unlikely in practice. Consequently, a standard way to tackle this problem is to rewrite Eq. (33) as a linear set of equations,

$$(\mathbf{S}^{-1} + \mathbf{N}^{-1})\,\mathbf{s}_{\rm WF} = \mathbf{N}^{-1}\,\mathbf{m}, \tag{34}$$

and solve these using some iterative method (e.g., Smith et al. 2007). Both CG and PCG techniques have been applied in this context and while the former was found to show a rather unsatisfactory convergence rate, it was demonstrated that this could be improved significantly albeit with the help of a rather advanced and involved (from the implementation point of view) preconditioner borrowed from multi-grid techniques (Smith et al. 2007).

The MF method originally proposed in this context by Elsner & Wandelt (2013) involves splitting the noise covariance into homogeneous and inhomogeneous parts by representing $\mathbf{N} \equiv \bar{\mathbf{N}} + \mathbf{T}$, where $\mathbf{T} = \tau\mathbf{I}$ is a homogeneous part, and $\tau = \min(\mathrm{diag}(\mathbf{N}))$. This leads to a split of the system matrix $\mathbf{S}^{-1} + \mathbf{N}^{-1}$ owing to the fact that,

$$\mathbf{N}^{-1} = \mathbf{T}^{-1} - \mathbf{T}^{-1}\,(\mathbf{N}^{-1} + \mathbf{T}^{-1})^{-1}\,\mathbf{T}^{-1}; \tag{35}$$

(see, e.g., Higham 2002 p.258). Subsequently, taking

$$\mathbf{C} \equiv \mathbf{S}^{-1} + \mathbf{T}^{-1}, \tag{36}$$
$$\mathbf{D} \equiv \mathbf{T}^{-1}\,(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1})^{-1}\,\mathbf{T}^{-1}, \tag{37}$$

and introducing the messenger field $\mathbf{t}$, Eq. (6), we can rewrite Eq. (33) in its messenger-field representation, that is,

$$\begin{aligned}(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1})\,\mathbf{t}^{(i+1)} &= \mathbf{T}^{-1}\,\mathbf{s}_{\rm WF}^{(i)} + \bar{\mathbf{N}}^{-1}\,\mathbf{m},\\ (\mathbf{S}^{-1} + \mathbf{T}^{-1})\,\mathbf{s}_{\rm WF}^{(i+1)} &= \mathbf{T}^{-1}\mathbf{t}^{(i+1)},\end{aligned} \tag{38}$$

with the former equation solved in the pixel and the latter in the harmonic domain and with the spherical harmonic transforms used to switch between these domains. These equations are equivalent to Eqs. (3) and (4) of Elsner & Wandelt (2013). Their numerical experiments showed that the solver tended to converge quickly to the solution given the desired precision and therefore the method was proposed as an efficient way to resolve the slow convergence problem of the CG method without the need for potentially complex preconditioners needed for an efficient PCG solver where both these methods should be applied directly to the initial problem, Eq. (34).

As argued earlier, Eqs. (38) are equivalent to a fixed-point iteration solver applied (see Eq. (5)),

$$\begin{aligned}&\left(\mathbf{I} - (\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1}\,\mathbf{T}^{-1}\,(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1})^{-1}\,\mathbf{T}^{-1}\right)\mathbf{s}_{\rm WF}\\ &\quad = (\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1}\,\mathbf{T}^{-1}\,(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1})^{-1}\,\bar{\mathbf{N}}^{-1}\,\mathbf{m},\end{aligned} \tag{39}$$

which can be rewritten in an explicitly iterative form as

$$\begin{aligned}(\mathbf{S}^{-1} + \mathbf{T}^{-1})\,\mathbf{s}_{\rm WF}^{(i+1)} = \mathbf{T}^{-1}\,(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1})^{-1}\,\mathbf{T}^{-1}\,\mathbf{s}_{\rm WF}^{(i)}\\ +\mathbf{T}^{-1}\,(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1})^{-1}\,\bar{\mathbf{N}}^{-1}\,\mathbf{m}.\end{aligned} \tag{40}$$

In the following section, we compare the performance of different solvers applied to Eqs. (34), (39), and (40). From the general consideration of the previous section our expectation is that the CG solver applied to Eq. (39), and equivalent to the PCG solution of Eq. (34) with $\mathbf{M} \equiv \mathbf{C} = \mathbf{S}^{-1} + \mathbf{T}^{-1}$, should perform better than the messenger-field solver, Eq. (40).

3.2. Simulated cases

To demonstrate and validate our analytical expectation we apply both these solvers to simulated data sets. These are obtained as follows. We first generate maps of three Stokes parameters, I, Q and U, in the Healpix pixelization (Górski et al. 2005) with the Healpix resolution parameter $n_{\rm side}$ set to 512. These maps are computed using a HEALpy routine, `synfast`, providing CMB power spectra as the input, as computed for the standard cosmological model with the current best values of the parameters (Planck Collaboration XIII 2016). In the following calculations we set the band-limit of the sky signal $\ell_{\rm max}$ to $2n_{\rm side}$. This is low enough to ensure the orthogonality of the relevant spherical harmonics over the grid of Healpix pixels. However, it leads to a rank-deficient signal covariance matrix. Consequently, hereafter, its inverse, $\mathbf{S}^{-1}$, is to be understood as a pseudo-inverse. We verified that in selected cases setting $\ell_{\rm max}$ to $3n_{\rm side}$ did not impact our conclusions. We add to these sky maps inhomogeneous, albeit uncorrelated noise with root mean

square (rms) changing over the sky as in the case of the WMAP observations[2].

We consider two cases with either full or partial sky observations. In this latter case, only 20% of the sky is observed corresponding to the polar cap regions as defined by the *Planck* HFI mask[3].

3.3. Numerical results

We consider the following solvers.

- CG applied to the redefined system, Eq. (39), which is equivalent to PCG applied to the original system Eq. (34) with a preconditioner given by $\mathbf{M} = \mathbf{S}^{-1} + \mathbf{T}^{-1}$ (in figures labeled "PCG");
- MF solver, Eq. (40) (in figures labeled "MF");
- MF method within three different cooling schemes as proposed in Elsner & Wandelt (2013), Huffenberger & Næss (2018), and Ramanah et al. (2017). In the first one, the value of the cooling parameter λ is adjusted adaptively after each iteration. This scheme requires an a priori knowledge (estimate) on the error measure (see Eq. (41) below) of the solution. For the purpose of the experiments, this is tightly approximated using the solution of the PCG solver. The scheme of Huffenberger & Næss (2018) defines a discrete grid of logarithmically spaced values of λ, which spans the range from 1 up to some suitable maximal value, which in our runs we set to $\lambda_{\rm max} = 10^4$. For each value of λ, a fixed number of iterations is performed. Though this scheme was suggested specifically for the map-making problem in order to avoid multiple time-consuming reads of the time-ordered data, for the sake of comparison we use it also for the WF experiments. Hereafter, we perform 10 iterations for each of the 16 values of λ and refer to this scheme as "16 × 10". For the case with partial sky observations, we continue with the fixed-point iterations (40) for $\lambda = 1$. The cooling scheme of (Ramanah et al. 2017, Algorithm 1) reduces λ by a constant factor, η, so $\lambda \rightarrow \lambda \times \eta$ and iterates as long as two consecutive approximations satisfy $\|\mathbf{s}^{(i)} - \mathbf{s}^{(i-1)}\|/\|\mathbf{s}^{(i)}\| < \epsilon$. In our experiments we start with $\lambda_{\rm max} = 10^4$, and we set $\eta = 3/4$ and $\epsilon \equiv 10^{-4}$.

We start the iterations with a vector of zeros as an initial guess.

The signal covariance, $\mathbf{S}$, is computed assuming the CMB power spectra as used for the simulations. The noise covariance is block-diagonal in the pixel domain but not proportional to $\mathbf{I}$ as the noise is assumed to be inhomogeneous. It is taken to be exactly the same as the noise covariance used for the simulations.

For all solvers, the inverse signal covariance is applied to a map-length pixel domain vector in the harmonic domain, when first the vector is represented by a vector of its harmonic coefficients computed with the help of a HEALpy routine, `map2alm`; these are subsequently weighted by the inverse of the power spectra and transformed back to the pixel domain using HEALpy's `alm2map` routine. The inverse noise covariance is applied to any pixel-domain vector directly with the elements corresponding to unobserved pixels set to zero.

We note that we always estimate the Wiener-filtered sky signal over the full sky. In all cases shown below we apply all the solvers to exactly the same input data sets.

We have validated our implementations by considering a simplified data set with white noise. In this case, the noise covariance $\mathbf{N}$ is proportional to the unit matrix, and the PCG solver with preconditioner $\mathbf{M} = \mathbf{S}^{-1} + \mathbf{T}^{-1}$, and the MF solver converged to within the numerical precision in a single step as expected, as in this case $\mathbf{T} = \mathbf{N}$. Moreover, in the cases of the actual simulated data sets used in our test, the results obtained with the different solvers are consistent.

3.3.1. Convergence metric

The Wiener-filter problem, Eq. (34), can be recast as a minimization of the functional

$$\chi^2(\mathbf{x}) = \mathbf{x}^t\mathbf{S}^{-1}\mathbf{x} + (\mathbf{m}-\mathbf{x})^t\mathbf{N}^{-1}(\mathbf{m}-\mathbf{x}). \tag{41}$$

Indeed, we have

$$\operatorname{argmin}\chi^2(\mathbf{x}) = (\mathbf{S}^{-1}+\mathbf{N}^{-1})^{-1}\mathbf{N}^{-1}\mathbf{m} = \mathbf{s}_{\rm WF}.$$

Algebraic manipulations show that χ^2 is directly related to the energy norm, as we have

$$\|\mathbf{s}_{\rm WF} - \mathbf{x}\|^2_{\mathbf{S}^{-1}+\mathbf{N}^{-1}} = \chi^2(\mathbf{x}) - \chi^2(\mathbf{s}_{\rm WF}).$$

We can therefore use physically motivated χ^2 as a convergence measure instead of the energy norm. We note that we expect that

$$\chi^2(\mathbf{x}) \rightarrow \chi^2(\mathbf{s}_{\rm WF}) = \mathbf{m}^t\,(\mathbf{S}+\mathbf{N})^{-1}\,\mathbf{m}. \tag{42}$$

As with the energy norm, Eq. (16), this asymptotic value of χ^2 cannot be straightforwardly computed without knowing the Wiener filter estimate precisely. However, we expect that it should be close to $\langle\chi^2(\mathbf{s}_{\rm WF})\rangle = n_{\rm Stokes}\,n_{\rm pix} \equiv n_{\rm DOF}$ within a small scatter on order of $O(\sqrt{2\,n_{\rm Stokes}\,n_{\rm pix}})$, if our assumptions about the sky signal and the noise are correct. Here the angle brackets denote the average over an ensemble of sky and noise realizations and $n_{\rm Stokes}$ stands for the number of considered Stokes parameters and is equal to 3 for most of our tests. We can therefore define the convergence in this case by requiring that the incremental change of χ^2 between consecutive iterations is not larger than some small fraction of $\langle\chi^2(\mathbf{s}_{\rm WF})\rangle$ (Elsner & Wandelt 2013). If the absolute value of the final χ^2 is statistically inconsistent with the expected one, this could be an indication of prematurely stalled convergence or of a problem with the model assumed for the measured data, $\mathbf{m}$.

Given the discussion of Sect. 2.3, we expect that in terms of minimizing the χ^2-measure, the PCG method with preconditioner $\mathbf{M} = \mathbf{S}^{-1} + \mathbf{T}^{-1}$ should be superior to the fixed-point iterations, Eq. (40).

In addition to the χ-measure we also plot the norm of the residual corresponding to the (preconditioned) problem as suggested in Ramanah et al. (2017). This is given by,

$$\mathbf{S}^{1/2}(\mathbf{S}^{-1}+\mathbf{N}^{-1})\,\mathbf{S}^{1/2}\mathbf{S}^{-1/2}\mathbf{s}_{\rm WF} = \mathbf{S}^{1/2}\mathbf{N}^{-1}\mathbf{m} \tag{43}$$

(Ramanah et al. 2017, Appendix C). This system is significantly better conditioned than the original one, Eq. (34). The corresponding relative norm of the residual then reads

$$\frac{\|\mathbf{S}^{1/2}\mathbf{N}^{-1}\mathbf{m} - \mathbf{S}^{1/2}(\mathbf{S}^{-1}+\mathbf{N}^{-1})\mathbf{x}\|}{\|\mathbf{S}^{1/2}\mathbf{N}^{-1}\mathbf{m}\|} = \frac{\|\mathbf{N}^{-1}\mathbf{m} - (\mathbf{S}^{-1}+\mathbf{N}^{-1})\mathbf{x}\|_{\mathbf{S}}}{\|\mathbf{N}^{-1}\mathbf{m}\|_{\mathbf{S}}}. \tag{44}$$

[2] We use specifically the noise pattern for the 9-year observation of the *V*-band, available from https://lambda.gsfc.nasa.gov/product/map/dr5/maps_band_r9_iqu_9yr_get.cfm.

[3] https://irsa.ipac.caltech.edu/data/Planck/release_2/ancillary-data/previews/HFI_Mask_GalPlane-apo2_2048_R2.00/index.html

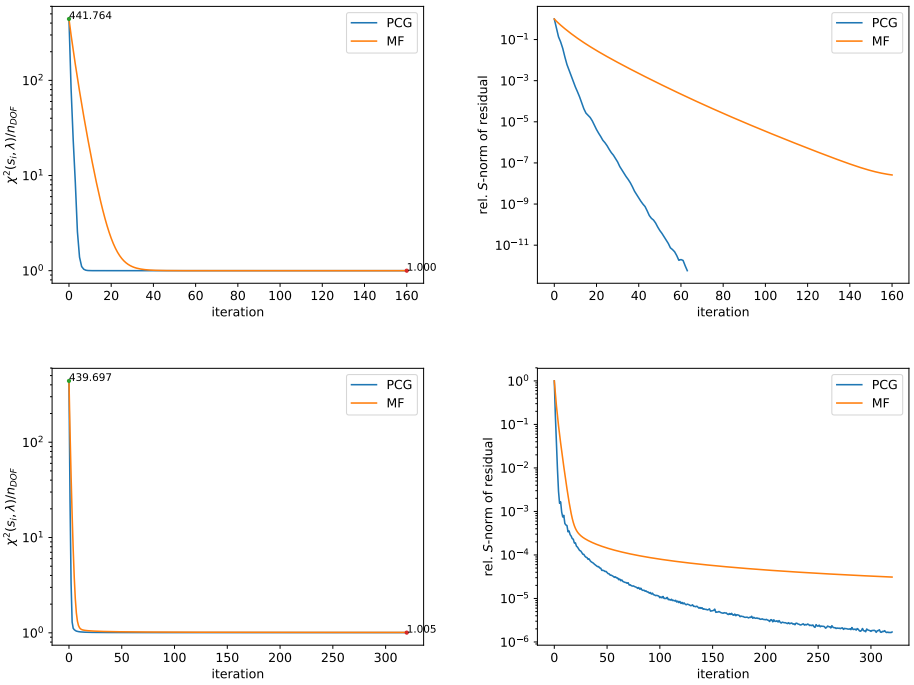

Fig. 1. Convergence of PCG and MF methods using two different convergence measures: χ^2 (Eq. (41), *left panels*), and the **S**-weighted relative residual (Eq. (44), *right panels*). The *top* and *bottom rows* show the case with the full and partial sky coverage, respectively.

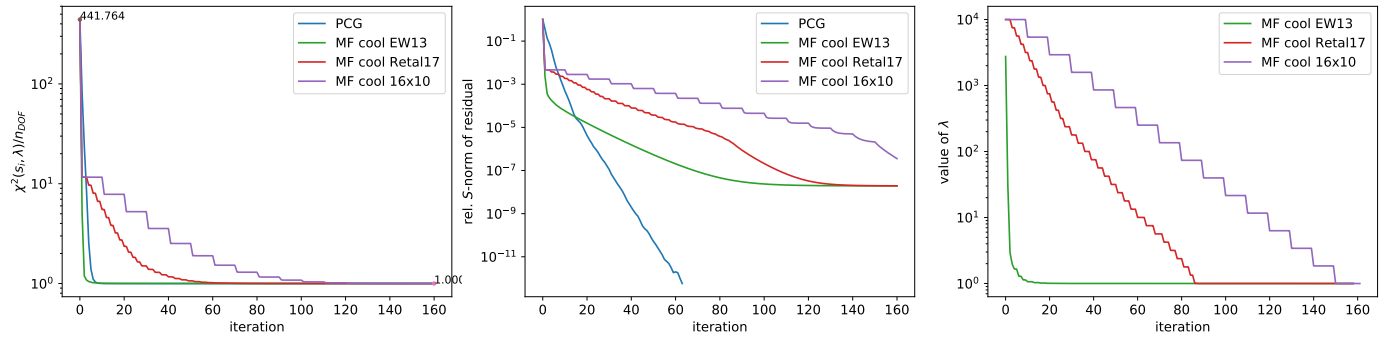

Fig. 2. Comparison of the convergence of the PCG solver and the MF technique with an adaptive cooling for different cooling prescriptions and using different convergence criteria: the χ^2, *left*, the **S**-weighted relative residual, middle. *Right panel*: values of λ as a function of the iteration as adapted by the different cooling schemes. These results are for the data sets with the full sky coverage.

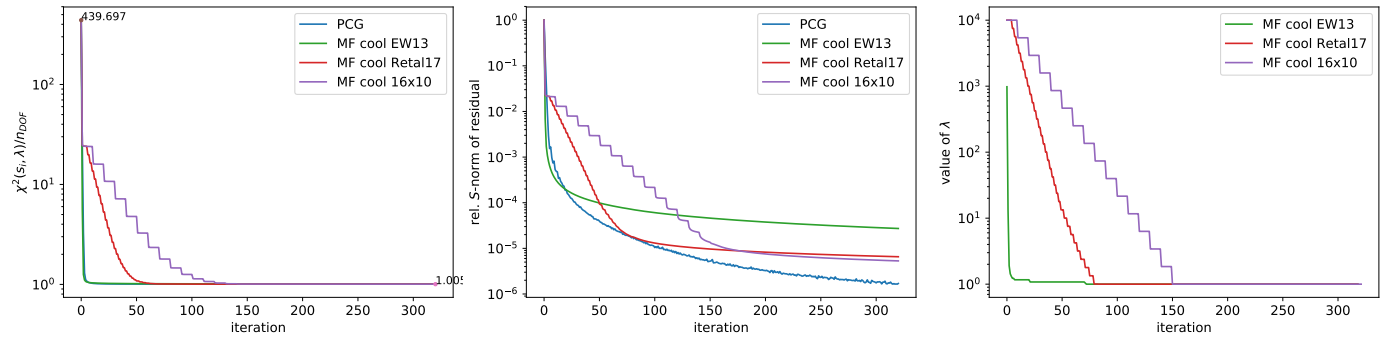

Fig. 3. As in Fig. 2 but for the data set with the partial sky coverage, $f_{\rm sky} = 0.2$.

3.3.2. Performance

Figure 1 shows a comparison between the PCG (Eq. (39)) and the MF (Eq. (40)) solvers as applied to the Wiener-filter problem. As expected, PCG indeed reduces the error significantly faster.

In Figs. 2 and 3, we compare PCG (with the MF preconditioner $\mathbf{S}^{-1} + \mathbf{T}^{-1}$ and $\lambda = 1$) with the MF method using the adaptive cooling schemes described above. We can see that PCG yields robust performance in all the test cases. In the case with full sky observations, the MF solvers (with or without cooling) reach their asymptotic convergence rate and exhibit a plateau of convergence on the level 10^{-7} of the relative **S**-norm of the residual. This is not the case for the PCG solver, which converges to the machine precision level. In the experiment with partial sky

coverage, we observe a decrease of the convergence rate for the PCG as well as the MF solvers due to significantly worse conditioning of the problem. However, even in this case, the PCG method is superior to the MF solvers. We expect that using more advanced preconditioners, which can alleviate the effect of very small eigenvalues, can bring a further significant improvement.

We note, however, that PCG appears to be outperformed by the MF method with the Elsner & Wandelt (2013) cooling proposal within the first ten or so iterations. As both methods solve different linear systems in this latter case, due to different values of λ, this does not contradict our conclusions in Sect. 2.3. This also does not change the overall assessment of the relative merits of both these techniques as no convergence is then ever reached in terms of any of the considered metrics. We discuss a possible origin of this behavior in Sect. 4.3, in the map-making context, where we suggest a simple antidote that could potentially further improve the performance of the PCG approach.

4. Application to map making

4.1. The problem

Data collected by modern, single-dish CMB experiments are modeled as

$$\mathbf{d} = \mathbf{P}\,\mathbf{m} + \mathbf{n}, \tag{45}$$

where $\mathbf{d}$ stands for a vector of all measurements, $\mathbf{m}$ is a pixelized map of the sky signal and $\mathbf{n}$ is the instrumental noise. We assume for simplicity that experimental beams are axially symmetric and that the sky signal $\mathbf{m}$ is already convolved with the beam. In this case, pointing matrix $\mathbf{P}$ simply defines which pixel of the map, $\mathbf{m}$, is observed at each measurement and with what weight it contributes to the measurement. In such cases, the pointing matrix is very sparse as it contains only one non-zero element per row for the total intensity measurements, or three for the polarization-sensitive ones. Moreover, $\mathbf{P}^t\,\mathbf{P}$ is either diagonal or block-diagonal with 3×3 blocks. If we assume that the instrumental noise is Gaussian with the covariance given by $\mathbf{N}$, a maximum-likelihood estimate of the sky signal can be written as

$$\mathbf{m}_{\rm ML} \equiv (\mathbf{P}^t\,\mathbf{N}^{-1}\,\mathbf{P})^{-1}\,\mathbf{P}^t\,\mathbf{N}^{-1}\,\mathbf{d}, \tag{46}$$

and therefore requires a solution of large linear system. The sizes of the involved object vary significantly, depending on the experiment, but the number of pixels in the map $\mathbf{m}$ can easily reach $O(10^6)$, while the number of measurements, $O(10^{12-15})$. Consequently, the system can only be solved iteratively, explicitly capitalizing on the structure and sparsity of the involved data objects.

Traditionally (e.g., de Gasperis et al. 2005; Cantalupo et al. 2010) the iterative method of choice was a PCG technique with a simple preconditioner given by

$$\mathbf{M} = \mathbf{P}^t\,{\rm diag}(\mathbf{N}^{-1})\,\mathbf{P}. \tag{47}$$

Hereafter we refer to this standard preconditioner as block-diagonal or Jacobi. However, more involved preconditioners have also been considered and found to be successful (e.g., Grigori et al. 2012; Szydlarski et al. 2014; Næss & Louis 2014; Puglisi et al. 2018).

More recently, Huffenberger & Næss (2018; see also Huffenberger 2018) proposed the application of the messenger-field technique to the map-making problem. Below, we discuss the approach of this former work. The proposal here is again to split the noise covariance into two parts, $\mathbf{N} \equiv \tilde{\mathbf{N}} + \mathbf{T}$, where $\mathbf{T} = \tau\mathbf{I}$, $\tau = \min({\rm diag}(\mathbf{N}))$. Subsequently using Eq. (35) we can rewrite the system matrix of the map-making equation, Eq. (46), as

$$\mathbf{P}^t\,\mathbf{N}^{-1}\,\mathbf{P} = \mathbf{P}^t\,\mathbf{T}^{-1}\,\mathbf{P} - \mathbf{P}^t\,\mathbf{T}^{-1}\,(\tilde{\mathbf{N}}^{-1} + \mathbf{T}^{-1})^{-1}\,\mathbf{T}^{-1}\,\mathbf{P}, \tag{48}$$

where the first term on the right-hand side corresponds to matrix $\mathbf{C}$ and the second one to matrix $\mathbf{D}$ as defined in Eq. (4). Following the formalism from Sect. 2.1 we can now write the messenger-field equations for this system, which read

$$(\tilde{\mathbf{N}}^{-1} + \mathbf{T}^{-1})\,\mathbf{t} = \tilde{\mathbf{N}}^{-1}\,\mathbf{d} + \mathbf{T}^{-1}\,\mathbf{P}\,\mathbf{m}_{\rm ML}, \tag{49}$$

$$(\mathbf{P}^t\,\mathbf{T}^{-1}\,\mathbf{P})\,\mathbf{m}_{\rm ML} = \mathbf{P}^t\,\mathbf{T}^{-1}\,\mathbf{t}, \tag{50}$$

with the messenger field $\mathbf{t}$ appearing explicitly, or

$$[\mathbf{I} - (\mathbf{P}^t\,\mathbf{T}^{-1}\,\mathbf{P})^{-1}\,\mathbf{P}^t\,\mathbf{T}^{-1}\,(\tilde{\mathbf{N}}^{-1} + \mathbf{T}^{-1})^{-1}\,\mathbf{T}^{-1}\,\mathbf{P}]\,\mathbf{m}_{\rm ML} = (\mathbf{P}^t\,\mathbf{T}^{-1}\,\mathbf{P})^{-1}\,\mathbf{P}^t\,\mathbf{N}^{-1}\,\mathbf{d}, \tag{51}$$

without it. We note that unlike in Eq. (6) the matrix $\mathbf{P}^t$ is taken out of the definition of the messenger field. Solving any of these two sets of equations using fixed-point iterations is equivalent to the messenger-field solver. For comparison we also solve the last equation using the CG technique. The latter is equivalent to solving the map-making equation, Eq. (46), using a PCG method with the preconditioner taken to be $\mathbf{M} = \mathbf{P}^t\,\mathbf{T}^{-1}\,\mathbf{P}$, which in the case under consideration is equivalent to the standard preconditioner.

We note, following Huffenberger & Næss (2018), that if the ith approximation issued by the fixed point method is unbiased, that is, $\langle \mathbf{x}^{(i)} - \mathbf{m} \rangle = 0$, where $\langle \dots \rangle$ denotes an average over noise realizations, then all the subsequent approximations will also be unbiased. In particular, if the initial guess is chosen to be an unbiased (e.g., simple binned, Eq. (61)) estimate of the sky signal, then all the following up estimates will be unbiased and the entire point of the iterations will be to converge on estimates with minimal statistical uncertainty. This is unlike the case of the PCG, where both statistical and systematic uncertainties are simultaneously improved on during the iterations.

4.2. Simulated data

We simulate mock time-ordered data $\mathbf{d}$ as a sum of two terms, one corresponding to the sky signal and the other to instrumental noise. These are computed as

$$\mathbf{d}_t = \mathbf{I}_{p(t)} + \mathbf{Q}_{p(t)}\,\cos 2\varphi(t) + \mathbf{U}_{p(t)}\,\sin 2\varphi(t) + \mathbf{n}_t, \tag{52}$$

where $p(t)$ denotes the sky pixels observed at time t and $\varphi(t)$ is the corresponding orientation of the polarizer. The signal terms are read off from the signal-only maps of Stokes parameters I, Q, and U, following the assumed scanning strategy defined by $p(t)$ and $\varphi(t)$. These maps are produced in the Healpix pixelization with the resolution parameter $n_{\rm side}$ set to 1024. These signals are random realizations of the CMB anisotropies corresponding to the currently preferred cosmological model (Planck Collaboration XIII 2016).

We produce three data sets with different statistical properties. Each data set comprises the data of a single detector that are however scaled to represent the performance of an entire detector array and thus are more representative of the current data. In all cases, we assume a raster scan pattern in the sky coordinates, when a rectangular sky patch made of 256×256 Healpix pixels is scanned either horizontally, that is, in right ascension, or vertically, in declination. The patch is centered at the equator. The

length of the simulated data vector is the same and roughly equal to 10^8.

In the first data set, the sky patch is first scanned horizontally and later vertically. The horizontal scanning assumes 256 complete sweeps (i.e., left-to-right followed by right-to-left), each pixel being sampled on average four times on each sweep. Once this is done, the declination is changed and the new horizontal scan commences. This is repeated 256 times with each horizontal scan corresponding to a different row of the Healpix pixels. The vertical scan is implemented in a similar way.

In this case we assume that the polarizer direction is quickly modulated so the full 2π angle is sampled within each single crossing of each sky pixel. This is ensured by setting the polarizer angle in the sky coordinates to follow a repeating sequence of $0, \pi/4, \pi/2, 3\pi/4$. In practice, this could mimic the case of an experiment using a smoothly rotating half wave plate.

For the second data set, we divide it into four equal consecutive subsets, each of which implements the same raster scan made of horizontal scanning within the first half of the subset followed by the vertical scan in the second half. However, the scanning is assumed to be faster and there is only one sample taken per pixel for each pixel crossing. This ensures the same data length. For each subset, the angle of the polarizer in the sky coordinates is fixed and equal to $0, \pi/4, \pi/2, 3\pi/4$. This scanning strategy mimics an experiment where the polarizer is stepped discretely only after each of the four subscans.

In the case of the third data set we progressively change the throw of the scan chop decreasing it gradually to half of the full scan width. We do so for both horizontal and vertical subscans. The scan speed is assumed fixed and tuned in a way that we obtain four observations of each pixel on a single pixel crossing. This produces a deeply observed core region where the number of observations per pixels can be as much as three orders of magnitude higher than the number of observations of the outer pixels. We also assume a smooth polarization angle rotation with the rotation speed fixed in such a way that the polarizer angle changes by 22.5° on a single pixel crossing.

This scan strategy is the most realistic from the three considered here reflecting the inhomogeneity of the sky sampling and allowing for imperfect sampling of the polarization angles per pixel. Consequently, for the scan parameters adopted in our simulations, the 3×3 blocks of the block-diagonal preconditioner display a range of condition numbers from 2 (perfect sampling) to over 20. The overall condition number of the block-diagonal preconditioner, which accounts for both sky and angle sampling inhomogeneities, is equal to $\sim 1.5 \times 10^4$.

We simulate the instrumental noise as a correlated noise with a power spectrum given by

$$\mathcal{P}(f) \equiv \sigma^2 \, (1 + \frac{f_{\rm knee}}{f}), \tag{53}$$

where $1/f_{\rm knee}$ is taken to be approximately 500 times longer than the sampling rate, corresponding to the length of a single full sweep of the sky. We further apodize the low frequency noise effectively flattening the noise spectrum for frequencies lower than a tenth of the knee frequency.

We consider two different noise levels, one ensuring a relatively high signal-to-noise ratio (S/N) of the resulting maps, with rms noise $\sigma^{Q,U}_{\rm rms} = \sigma^{I}_{\rm rms} \cdot \sqrt{2} \simeq 2\,\mu$K for the recovered Q and U maps, and the other with lower S/N, corresponding to $\sigma^{Q,U}_{\rm rms} \simeq 30\,\mu$K. We refer to these cases as the low- and high-noise data.

We note, that if the instrumental noise were white then the two first scanning strategies would have been equivalent and the standard, block-diagonal preconditioners in both these cases would have been identical. This is however not the case in the presence of the correlated noise. In fact we expect that the off-diagonal noise correlation of the recovered Q and U maps should be small for the first data set with the quickly rotating HWP while they should be non-negligible in the case of the second data set with the stepped polarizer, leading to different convergence patterns of the studied solvers.

4.3. Numerical results

4.3.1. Convergence metric

For measuring the error of an approximation $\mathbf{x}$ we consider, following Huffenberger & Næss (2018), a χ-measure

$$\chi^2(\mathbf{x}) \;=\; (\mathbf{d} - \mathbf{P}\mathbf{x})^t \, \mathbf{N}^{-1} (\mathbf{d} - \mathbf{P}\mathbf{x}). \tag{54}$$

Analogously to the Wiener-filtering application, this measure is minimized by the maximum-likelihood estimate (46) and is equivalent to the energy norm of the error, Eq. (16), with respect to the system matrix, $\mathbf{A} \equiv \mathbf{P}^t \, \mathbf{N}^{-1} \, \mathbf{P}$. Indeed,

$$\begin{aligned} \chi^2(\mathbf{x}) \;&=\; \|\mathbf{x} - \mathbf{m}_{\rm ML}\|_{\mathbf{P}^t \mathbf{N}^{-1} \mathbf{P}} \\ &+\; \mathbf{d}^t \, (\mathbf{N}^{-1} - \mathbf{N}^{-1}\mathbf{P} \, (\mathbf{P}^t \, \mathbf{N}^{-1} \, \mathbf{P})^{-1} \, \mathbf{P}^t \, \mathbf{N}^{-1}) \, \mathbf{d}, \end{aligned} \tag{55}$$

and thus,

$$\chi^2(\mathbf{m}_{\rm ML}) \;=\; \mathbf{d}^t \, (\mathbf{N}^{-1} - \mathbf{N}^{-1}\mathbf{P} \, (\mathbf{P}^t \, \mathbf{N}^{-1} \, \mathbf{P})^{-1} \, \mathbf{P}^t \, \mathbf{N}^{-1}) \, \mathbf{d}. \tag{56}$$

As before this value is not directly available. However, we can compute the average value of $\chi^2(\mathbf{m}_{\rm ML})$ over the statistical ensemble of the input data realizations and use it as a benchmark for the convergence using the χ^2-measure. This can be done analytically, observing that the matrix on the right-hand side of Eq. (56) is a projection operator, which projects out all time-domain modes, which are sky stationary, that is, they are objects of the form $\mathbf{P}\,\mathbf{y}$ for some arbitrary pixel-domain object $\mathbf{y}$. If so,

$$\chi^2(\mathbf{m}_{\rm ML}) \;=\; \mathbf{n}^t \, (\mathbf{N}^{-1} - \mathbf{N}^{-1}\mathbf{P} \, (\mathbf{P}^t \, \mathbf{N}^{-1} \, \mathbf{P})^{-1} \, \mathbf{P}^t \, \mathbf{N}^{-1}) \, \mathbf{n}, \tag{57}$$

and

$$\begin{aligned} \langle \chi^2(\mathbf{m}_{\rm ML}) \rangle_{\rm noise} \;&=\; {\rm tr} \, (\mathbf{I} - \mathbf{N}^{-1/2}\mathbf{P} \, (\mathbf{P}^t \, \mathbf{N}^{-1} \, \mathbf{P})^{-1} \, \mathbf{P}^t \, \mathbf{N}^{-1/2}) \\ &=\; n_t - n_{\rm Stokes} \, n_{\rm pix} \;\equiv\; n_{\rm DOF}, \end{aligned} \tag{58}$$

where n_t and $n_{\rm pix}$ denote the sizes of the data set in time and pixel domains, respectively, and assuming that the system matrix, $\mathbf{P}^t \, \mathbf{N}^{-1} \, \mathbf{P}$, is non-singular, and considering that $\mathbf{N}^{-1/2} \, \langle \mathbf{n} \, \mathbf{n}^t \rangle \, \mathbf{N}^{-1/2} = \mathbf{I}$. With this value, we then define the convergence criterion in terms of the χ^2-measure by requiring that the incremental change of χ^2 between two consecutive iterations is sufficiently small as compared to $n_{\rm DOF}$.

We note that in the figures, in order to make the behavior of the χ-measure more conspicuous, instead of the χ^2 itself, we plot its relative difference with respect to the minimal value of χ^2 derived within the PCG iterations, which we denote as $\chi^2(\mathbf{m}_{\rm min})$. The plotted quantity is then given by,

$$\chi^2(\mathbf{x}) / \chi^2(\mathbf{m}_{\rm min}) - 1. \tag{59}$$

For completeness we also plot the standard relative residual defined as

$$\frac{\|\mathbf{P}^t \, \mathbf{N}^{-1} \mathbf{d} - (\mathbf{P}^t \, \mathbf{N}^{-1} \, \mathbf{P}) \, \mathbf{x}\|}{\|\mathbf{P}^t \, \mathbf{N}^{-1} \mathbf{d}\|}. \tag{60}$$

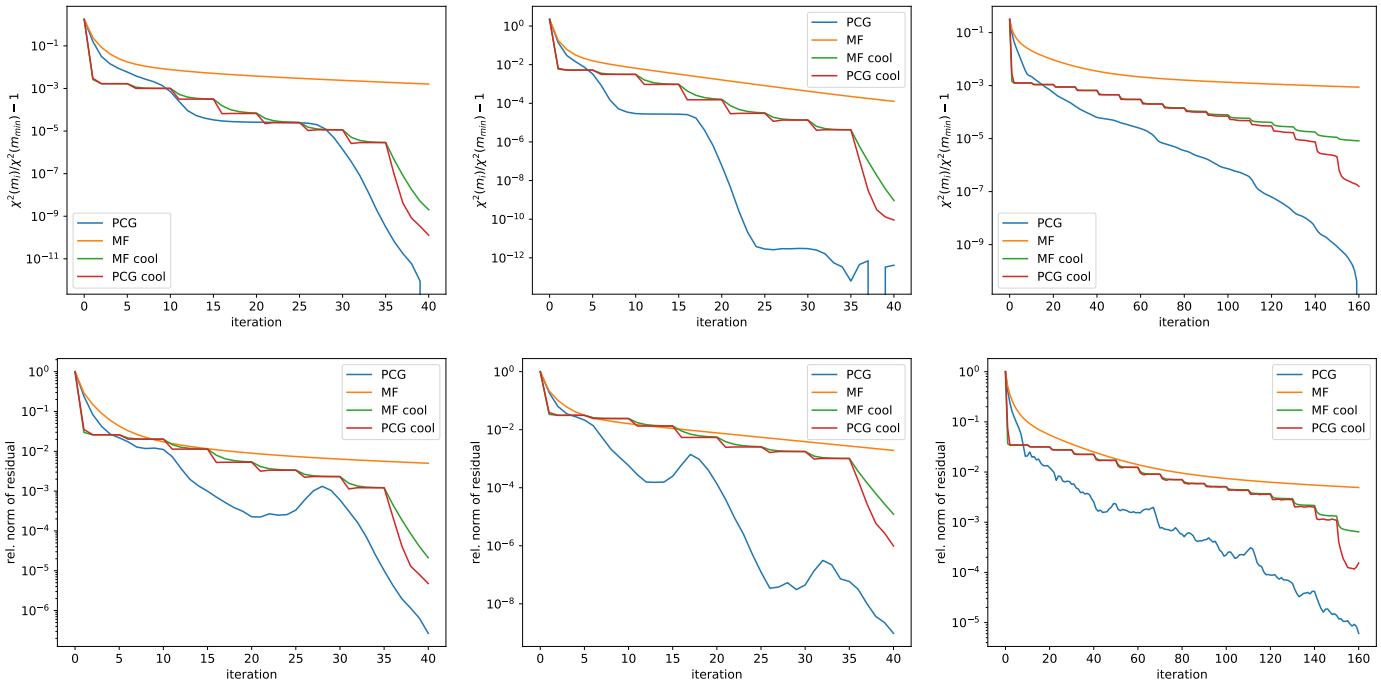

Fig. 4. Comparison of the convergence for the PCG, the messenger-field methods stand-alone and incorporated within a cooling scheme, for the first, second, and third simulated data sets (*from left to right panels*, respectively), assuming a low noise level. The cooling scheme is 8×5 for the first and second data sets and 16×10 for the third.

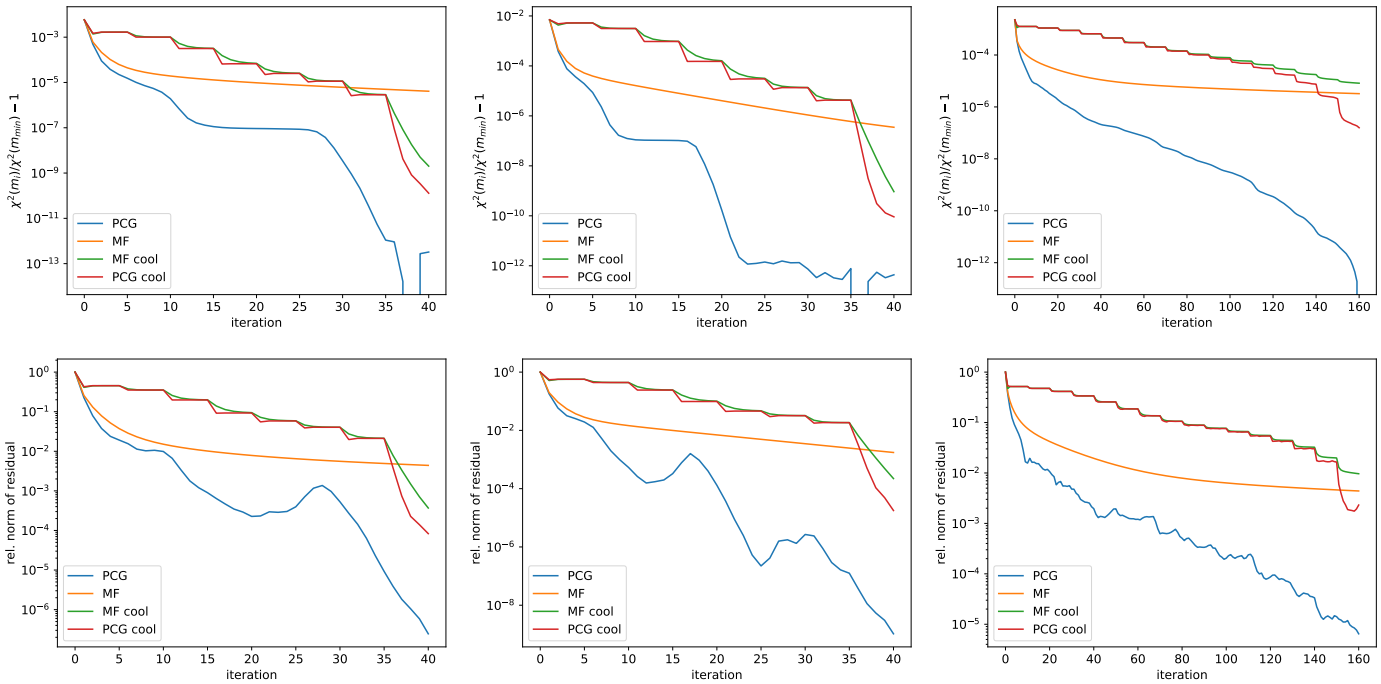

Fig. 5. As in Fig. 4 but for the high noise level.

4.3.2. Performance

We reconstruct the sky signal from the data using different solvers as discussed here and compare their relative performances. In all cases, we use the same pixelization for the recovered maps as we used for the simulations, that is, Healpix with $n_{\rm side} = 1024$. We validated our implementation by running cases with noise-free data and recovering the input maps within the numerical precision. We also found that the results produced by different solvers for each of the data sets agree.

We first compare the convergence of the PCG solver, of the MF iterations without cooling, and of the PCG and MF with the 8×5 cooling scheme. The results for the low-noise and high-noise cases are given separately in Figs. 4 and 5.

We can see that as claimed in Huffenberger & Næss (2018) the MF with cooling technique indeed reaches higher accuracy

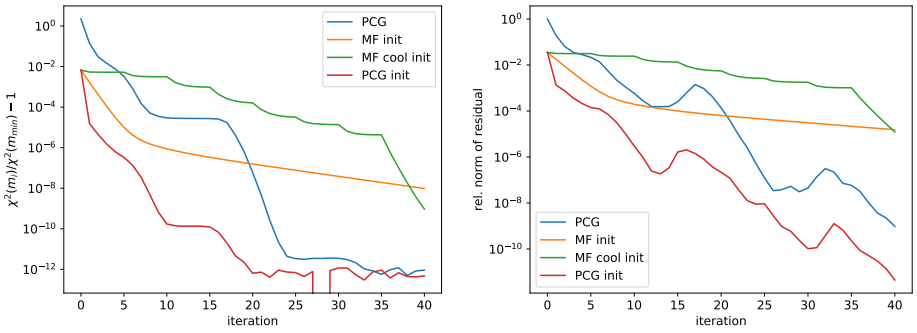

Fig. 6. Comparison of the convergence rates of different iterative solvers for a non-zero starting vector, $\mathbf{m}^{(0)}$, as given in Eq. (61). For comparison the blue curve shows the case of the PCG with a starting vector of zeros. The results are for the second scanning strategy and the low-noise case.

in comparison to MF without cooling. However, the standard PCG is in these experiments still superior. As in the Wiener-filter application (Sect. 3.3), we observe that in the low-noise cases the cooling technique, used either with PCG or with MF, improves more rapidly on the solution within the first iterations than the PCG method with no cooling. We attribute this to the fact that during those initial iterations the cooling method solves a modified system of the initial equations with an assumed large value of λ. The approximate solution derived on these first iterations is then equivalent to a simple binned map. This for the low noise cases provides a good rendition of the sky signal, thus leading to an abrupt decrease of the residuals. In the absence of cooling the PCG technique initiated with vectors of zero needs to perform at least a few iterations to reach a comparably good solution. We can however improve on the performance of the stand-alone PCG by using a simple binned map, given by

$$\mathbf{m}^{(0)} = (\mathbf{P}^t \,\mathrm{diag}(\mathbf{N})^{-1}\mathbf{P})^{-1}\mathbf{P}^t \,\mathrm{diag}(\mathbf{N})^{-1}\mathbf{d}, \tag{61}$$

as the starting vector for the PCG solver. Such a map is quickly computable and thus can be readily available at the onset of the solution.

We illustrate these considerations in Fig. 6, where we compare the convergence of the PCG run with the initial vector made of zeros and the convergence of the solvers: stand-alone PCG, MF with cooling, and PCG with cooling, assuming $\mathbf{m}^{(0)}$ (Eq. (61)) as the initial vector. The results shown are for the low-noise case. As expected there is a significant improvement in the overall performance of the PCG method relative to the other solvers but also as compared to the case of the vanishing initial guess. This showcases the importance of the appropriate choice of the initial guess for the PCG approaches in the cases of high-S/N solutions. As the new CMB data sets target predominantly CMB B-mode polarization, the maps of Stokes parameters will increasingly have very high S/N and this observation may be therefore of importance for their data analysis.

We note that all numerical experiments considered here involve non-singular linear systems of equations. If singularities are present then both PCG and MF solvers will typically saturate before reaching the convergence. For the cooled MF this may however not be the case. In particular, if the modified linear systems with $\lambda \gtrsim 1$ are singularity free then λ is effectively a regularization parameter. In such cases the cooled MF may reach the residual level better than the other methods thanks to its ability to adapt amplitudes of the singular modes present in the solution. This however does not change the fact that if these modes are truly singular then their true amplitudes cannot ever be recovered. If the regularization is the appropriate approach to adapt in a given application, then this could also be done in the case of the other solvers.

5. Conclusions

We have shown that the messenger-field solvers of sets of linear equations perform fixed-point iterations of an appropriately preconditioned system of equations. Consequently, in general they are expected to display inferior performance to that of a conjugate gradient solver applied to the same preconditioned systems or, equivalently, to that of a PCG solver with the same preconditioner as implicitly used in the messenger-field method in order to precondition the initial problem. We have backed up this contention with analytic arguments and demonstrated it using numerical experiments involving two applications drawn from modern CMB data analysis practice: Wiener filters and the CMB map-making problem. In addition to the basic implementations of the MF method (Sect. 2.1), we have considered MF solvers combined with the cooling technique (Sect. 2.2, Elsner & Wandelt 2013; Huffenberger & Næss 2018), and have shown via numerical results that the cooled MF methods with the cooling schemes as proposed in the literature outperform the standard MF approach. However, the PCG solvers with the preconditioner motivated by the MF methods tend to reach convergence the quickest.

We have compared the performance of the studied methods from the perspective of the number of iterations needed to reach convergence. However, our conclusions are expected to also be directly applicable to considerations involving time-to-solution, as the computational cost per iteration incurred in the different methods is found to be roughly comparable.

We therefore conclude that looking towards the future, advanced preconditioning coupled with the conjugate gradient technique offers the most promise as an expeditious solver, ahead of the messenger-field approach. While at this time, the PCG solvers, with the standard block-diagonal preconditioner (Eq. (47)) in the map-making case, and the preconditioner given by $\mathbf{S}^{-1} + \mathbf{T}^{-1}$ (Eq. (36)) in the Wiener filter case, with a potentially appropriately adapted initial guess, should outperform the currently proposed messenger-field approaches. We also note that better preconditioners have already been proposed in particular in the map-making context (e.g., Grigori et al. 2012; Szydlarski et al. 2014). This notwithstanding, the messenger-field approach may be found of interest in some specific applications.

In the context of the PCG methods, we have found that the convergence may be sped up by an appropriate choice of initial vector. While the gain is largely negligible for the cases with a low-S/N solution, it can become significant if the solution is expected to have high-S/N content. We have found this effect particularly relevant for the map-making procedure, where we have shown that the choice of the simple binned map as the initial vector can result in a significant improvement of the map-making solver convergence.

Acknowledgements. We thank Dominic Beck and Josquin Errard for their help with simulations and insightful discussions and Kevin Huffenberger and Sigurd Næss for their comments on the manuscript. We acknowledge use of HEALpy. The first two authors' work was supported by the NLAFET project as part of European Union's Horizon 2020 research and innovation program under grant 671633. RS acknowledges support of the French National Research Agency (ANR) contract ANR-17-C23-0002-01 (project B3DCMB). This research used resources of the National Energy Research Scientific Computing Center (NERSC), a DOE Office of Science User Facility supported by the Office of Science of the US Department of Energy under Contract No. DE-AC02-05CH11231.

References

Björck, A. K. 2015, Numerical Methods in Matrix Computations (Cham: Springer), XVI, 800
Bunn, E. F., Fisher, K. B., Hoffman, Y., et al. 1994, ApJ, 432, L75
Cantalupo, C. M., Borrill, J. D., Jaffe, A. H., Kisner, T. S., & Stompor, R. 2010, ApJS, 187, 212
de Gasperis, G., Balbi, A., Cabella, P., Natoli, P., & Vittorio, N. 2005, A&A, 436, 1159
Elsner, F., & Wandelt, B. D. 2013, A&A, 549, A111
Golub, G. H., & Van Loan, C. F. 1996, in Matrix Computations, 3rd edn. (Baltimore, MD: Johns Hopkins University Press), XXX, 698
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Grigori, L., Stompor, R., & Szydlarski, M. 2012, Proc. Int. Conf. on High Performance Computing, Networking, Storage and Analysis, SC '12 (Los Alamitos, CA, USA: IEEE Computer Society Press), 91:1
Higham, N. J. 2002, Accuracy and Stability of Numerical Algorithms, 2nd edn. (Philadelphia, PA: SIAM), XXX, 680
Horn, R. A. 2013, Matrix Analysis, 2nd edn. (Cambridge: Cambridge University Press), XVIII, 643
Huffenberger, K. M. 2018, MNRAS, 476, 3425
Huffenberger, K. M., & Næss, S. K. 2018, ApJ, 852, 92
Janssen, M. A., & Gulkis, S. 1992, NATO Advanced Science Institutes (ASI) Ser. C, 359, 391
Liesen, J., & Strakoš, Z. 2013, Krylov Subspace Methods: Principles and Analysis, Numerical Mathematics and Scientific Computation (Oxford: Oxford University Press), XVI, 391
Næss, S. K., & Louis, T. 2014, JCAP, 8, 045
Planck Collaboration XIII. 2016, A&A, 594, A13
Poletti, D., Fabbian, G., Le Jeune, M., et al. 2017, A&A, 600, A60
Puglisi, G., Poletti, D., Fabbian, G., et al. 2018, A&A, 618, A62
Ramanah, D., Lavaux, G., & Wandelt, B. 2017, MNRAS, 468, 1782
Saad, Y. 2003, Iterative Methods for Sparse Linear Systems, 2nd edn. (Philadelphia, PA: SIAM), XVIII, 528
Seljebotn, D. S., Bærland, T., Eriksen, H. K., Mardal, K. A., & Wehus, I. K. 2017, A&A, submitted, [arXiv:1710.00621]
Smith, K. M., Zahn, O., & Doré, O. 2007, Phys. Rev. D, 76, 043510
Strakoš, Z., & Tichý, P. 2005, BIT, 45, 789
Szydlarski, M., Grigori, L., & Stompor, R. 2014, A&A, 572, A39

Appendix A: Implementation of PCG for Wiener filter

In the context of solving the Wiener-filter problem (Eq. (33)), each step of the fixed-point method (Eq. (12), resp. Eq. (40)) requires one direct and one inverse spherical harmonic transforms, which are assumed to be the most time-consuming elements of the solution process. To keep the same number of transforms in each PCG iteration, we cannot apply a first matrix, $\mathbf{A} = \mathbf{S}^{-1} + \mathbf{N}^{-1}$, and then precondition the residual by $\mathbf{C}^{-1} \equiv (\mathbf{S}^{-1}+\mathbf{T}^{-1})^{-1}$ as is done in one of the standard PCG implementations listed below in Algorithm A.1. This implementation involves two direct and two inverse transforms: in the evaluation of $\mathbf{A}\mathbf{p}^{(i-1)}$ and in $\mathbf{C}^{-1}\mathbf{r}^{(i)}$.

Algorithm A.1 PCG for $\mathbf{A}\mathbf{s} = \mathbf{b}$ with the preconditioner $\mathbf{C}$

Given $\mathbf{s}^{(0)}$, $\mathbf{r}^{(0)} = \mathbf{b} - \mathbf{A}\mathbf{s}^{(0)}$, $\widetilde{\mathbf{r}}^{(0)} = \mathbf{C}^{-1}\mathbf{r}^{(0)}$, $\mathbf{p}^{(0)} = \widetilde{\mathbf{r}}^{(0)}$.
For $i = 1, 2, \ldots$

$$
\begin{aligned}
\gamma^{(i-1)} &= \frac{(\mathbf{r}^{(i-1)}, \widetilde{\mathbf{r}}^{(i-1)})}{(\mathbf{p}^{(i-1)}, \mathbf{A}\mathbf{p}^{(i-1)})},\\
\mathbf{s}^{(i)} &= \mathbf{s}^{(i-1)} + \gamma^{(i-1)}\mathbf{p}^{(i-1)}\,,\\
\mathbf{r}^{(i)} &= \mathbf{r}^{(i-1)} - \gamma^{(i-1)}\mathbf{A}\mathbf{p}^{(i-1)}\,,\\
\widetilde{\mathbf{r}}^{(i)} &= \mathbf{C}^{-1}\mathbf{r}^{(i)}\,,\\
\delta^{(i)} &= \frac{(\mathbf{r}^{(i)}, \widetilde{\mathbf{r}}^{(i)})}{(\mathbf{r}^{(i-1)}, \widetilde{\mathbf{r}}^{(i-1)})},\\
\mathbf{p}^{(i)} &= \widetilde{\mathbf{r}}^{(i)} + \delta^{(i)}\mathbf{p}^{(i-1)}\,.
\end{aligned}
$$

Using the formula for $\mathbf{p}^{(i)}$, $\widetilde{\mathbf{r}}^{(i)}$ and the form of the matrix $\mathbf{A} = \mathbf{C} - \mathbf{D}$, we can write

$$
\begin{aligned}
\mathbf{A}\mathbf{p}^{(i)} &= \mathbf{A}(\widetilde{\mathbf{r}}^{(i)} + \delta^{(i)}\mathbf{p}^{(i-1)})\\
&= \mathbf{C}\mathbf{C}^{-1}\mathbf{r}^{(i)} - \mathbf{D}\widetilde{\mathbf{r}}^{(i)} + \delta^{(i)}\mathbf{A}\mathbf{p}^{(i-1)}\\
&= \mathbf{r}^{(i)} - \mathbf{D}\widetilde{\mathbf{r}}^{(i)} + \delta^{(i)}\mathbf{A}\mathbf{p}^{(i-1)}.
\end{aligned}
$$

Therefore the vector $\mathbf{A}\mathbf{p}^{(i)}$ can be computed recursively without spherical harmonic transforms and the cost of one PCG iteration is the same (in terms of spherical harmonic transforms) as the cost of one iteration of the fixed-point method, Eq. (12).

Another formula that proved in our numerical experiments to be more stable (yet slightly more costly) is to simultaneously evaluate the vectors $\mathbf{C}^{-1}\mathbf{r}^{(i)}$ and $\mathbf{S}^{-1}\mathbf{C}^{-1}\mathbf{r}^{(i)}$ (recall that $\mathbf{A}\mathbf{C}^{-1}\mathbf{r}^{(i)} = (\mathbf{S}^{-1} + \mathbf{N}^{-1})\mathbf{C}^{-1}\mathbf{r}^{(i)}$). This can be done using direct spherical harmonic transform of one vector, $\mathbf{r}^{(i)}$, and inverse spherical harmonic transform of two vectors[4]. We then simply update $\mathbf{A}\mathbf{p}^{(i)} = \mathbf{A}\mathbf{C}^{-1}\mathbf{r}^{(i)} + \delta^{(i)}\mathbf{A}\mathbf{p}^{(i-1)}$.

Moreover, the properties of (P)CG also allow to evaluate the decrease of the χ-measure without computing it explicitly using Eq. (41) (this computation involves $\mathbf{S}^{-1}$ and therefore also direct and inverse spherical harmonic transforms). The evaluation proposed below is numerically stable; see a thorough analysis in Strakoš & Tichý (2005). There holds

$$
\|\mathbf{s} - \mathbf{s}^{(0)}\|^2_{\mathbf{A}} = \sum_{j=1}^{i} \gamma_j\,(\mathbf{r}^{(j)}, \widetilde{\mathbf{r}}^{(j)}) + \|\mathbf{s} - \mathbf{s}^{(i)}\|^2_{\mathbf{A}}.
$$

[4] With the typical computational cost significantly smaller than twice the cost of one single inverse spherical harmonic transform.

Using the above discussion on the relationship between the energy norm and the χ-measure, we have

$$
\chi^2(\mathbf{s}^{(0)}) = \sum_{j=1}^{i} \gamma_j\,(\mathbf{r}^{(j)}, \widetilde{\mathbf{r}}^{(j)}) + \chi^2(\mathbf{s}^{(i)}).
$$

After computing $\chi^2(\mathbf{s}^{(0)})$ (that is for zero initial approximation $\mathbf{s}^{(0)} = 0$ equal to $\mathbf{m}^t\mathbf{N}^{-1}\mathbf{m}$) we can therefore simply evaluate the χ-measure in every PCG iteration using already computed scalar quantities without any additional spherical harmonic transforms.

Appendix B: Proof of convergence of the messenger-field method

In this appendix we prove that the messenger-field method for Wiener filter is (asymptotically) converging. Following the discussion in Sect. 2.3, we prove the convergence by showing that the eigenvalues of $\mathbf{C}^{-1}\mathbf{D}$ are, in the absolute value, smaller than unity. First, we note that, since $\mathbf{T} = \tau\mathbf{I}$,

$$
\mathbf{C}^{-1}\mathbf{D} = [(\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1}\mathbf{T}^{-1}][(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1})^{-1}\mathbf{T}^{-1}]
$$

is given by multiplication of two symmetric matrices (in brackets).

Algebraic manipulations then yield

$$
\begin{aligned}
(\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1}\mathbf{T}^{-1} &= \mathbf{S}\,(\mathbf{S} + \mathbf{T})^{-1},\\
(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1})^{-1}\mathbf{T}^{-1} &= \bar{\mathbf{N}}\,(\bar{\mathbf{N}} + \mathbf{T})^{-1};
\end{aligned}
$$

see also (Elsner & Wandelt 2013, Eq. (5)). Since $\mathbf{S}$, $\bar{\mathbf{N}}$ are symmetric positive semidefinite and $\tau > 0$, the eigenvalues of the matrices above are in the interval $[0, 1)$. For a symmetric matrix $\mathbf{B}$ there holds $\|\mathbf{B}\|_2 = \rho(\mathbf{B})$.

Finally,

$$
\rho(\mathbf{C}^{-1}\mathbf{D}) \le \|\mathbf{C}^{-1}\mathbf{D}\|_2 \le \|\mathbf{S}(\mathbf{S} + \mathbf{T})^{-1}\|_2 \cdot \|\bar{\mathbf{N}}(\bar{\mathbf{N}} + \mathbf{T})^{-1}\|_2 < 1.
$$

In order to present a close relationship of the derivation of the messenger-field with the Schur complement methods, we present an alternative proof below. An analogous derivation can be used also for proving the convergence of MF in the map-making application.

We start by rewriting Eq. (38) as the system

$$
\underbrace{\begin{pmatrix} \bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1} & -\mathbf{T}^{-1} \\ -\mathbf{T}^{-1} & \mathbf{S}^{-1} + \mathbf{T}^{-1} \end{pmatrix}}_{\equiv\,\mathbb{A}.} \begin{pmatrix} \mathbf{t} \\ \mathbf{s}_{\mathrm{WF}} \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{N}}^{-1}\,\mathbf{m} \\ 0 \end{pmatrix}
$$

The reduced system (after the elimination of the messenger field $\mathbf{t}$) then corresponds to forming the Schur complement $\mathbb{S}$ of $\mathbb{A}$,

$$
\mathbb{S} \equiv (\mathbf{S}^{-1} + \mathbf{T}^{-1}) - \mathbf{T}^{-1}(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1})^{-1}\mathbf{T}^{-1},
$$

and solving

$$
\mathbb{S}\mathbf{s}_{\mathrm{WF}} = \mathbf{T}^{-1}(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1})^{-1}\bar{\mathbf{N}}^{-1}\,\mathbf{m}.
$$

The MF iterations are obtained by multiplying (preconditioning) the above system by $(\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1}$ from the left.

Now we show the bounds on the eigenvalues of

$$
(\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1}\mathbf{T}^{-1}(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1})^{-1}\mathbf{T}^{-1} = \mathbf{I} - (\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1}\mathbb{S}.
$$

Since $\mathbb{A}$ is an SPD matrix, its Schur complement $\mathbb{S}$ is also SPD. Moreover, the spectrum of $(\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1}\mathbb{S}$ satisfies

$$\Lambda\left((\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1}\mathbb{S}\right) = \Lambda\left((\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1/2}\,\mathbb{S}\,(\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1/2}\right),$$

and therefore the eigenvalues of $(\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1}\mathbb{S}$ are positive. Plugging into the above equation the formula for the Schur complement $\mathbb{S}$, we have

$$\begin{aligned}
&\Lambda\left((\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1}\mathbb{S}\right) \\
&\quad = \Lambda\left(\mathbf{I} - (\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1/2}(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1})^{-1}(\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1/2}\right) \\
&\quad = 1 - \Lambda\left((\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1/2}(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1})^{-1}(\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1/2}\right).
\end{aligned}$$

The matrix $(\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1/2}(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1})^{-1}(\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1/2}$ is symmetric positive semidefinite. Altogether,

$$\Lambda\left((\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1}\mathbb{S}\right) \in (0, 1].$$

Consequently,

$$\Lambda\left(\mathbf{C}^{-1}\mathbf{D}\right) = 1 - \Lambda\left((\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1}\mathbb{S}\right) \in [0, 1).$$

Finally, we show that the asymptotic convergence of the error in the Euclidean norm $\|\epsilon^{(i)}\|$, which is assured by the fact that $\rho(\mathbf{C}^{-1}\mathbf{D}) < 1$ (see, e.g., Saad 2003, Sect. 4.2), proves the asymptotic convergence of the error $\epsilon^{(i)}$ in any norm $\|\cdot\|_*$. Here we use the equivalence of norms on finite-dimensional spaces; see, e.g., Horn (2013, Corollary 5.4.6 and Definition 5.4.7). In particular, given any norm $\|\cdot\|_*$, there exist positive constants c_*, C_*, such that

$$c_*\|\mathbf{v}\|_* \le \|\mathbf{v}\| \le C_*\|\mathbf{v}\|_*, \quad \text{for all} \quad \mathbf{v}.$$

Consequently,

$$\|\epsilon^{(i)}\|_* \le (c_*)^{-1}\,\|\epsilon^{(i)}\| \to 0, \quad \text{for} \quad i \to \infty.$$