Principles of Systematic Upscaling

Achi Brandt

1 Introduction

Despite their dizzying speed, modern supercomputers are still incapable of handling many most vital scientific problems. This is primarily due to *the scale gap*, which exists between the microscopic scale at which physical laws are given and the much larger scale of phenomena we wish to understand.

This gap implies, first of all, a huge number of *variables* (e.g., atoms or gridpoints or pixels), and possibly even a much larger number of *interactions* (e.g., one force between every pair of atoms). Moreover, computers simulate physical systems by moving *few variables at a time*; each such move must be extremely small, since a larger move would have to take into account all the motions that should in parallel be performed by all other variables. Such a computer simulation is particularly incapable of moving the system across large-scale *energy barriers*, which can each be crossed only by a large, *coherent* motion of very many variables.

This type of obstacles makes it impossible, for example, to calculate the properties of nature's building blocks (elementary particles, atomic nuclei, etc.), or to computerize chemistry and materials science, so as to enable the design of materials, drugs and processes, with enormous potential benefits for medicine, biotechnology, nanotechnology, agriculture, materials science, industrial processing, etc. With current common methods the amount of computer processing often increases so steeply with problem size, that even much faster computers will not do.

Past studies have demonstrated that scale-born slownesses can often be overcome by multiscale algorithms. Such algorithms have first been developed in the form of fast *multigrid solvers* for discretized PDEs [1], [2], [3], [4], [13], [15], [14]. These solvers are based on two processes: (1) classical *relaxation* schemes, which are generally slow to converge but fast to *smooth* the error function; (2) approximating the smooth error on a *coarser grid* (typically having twice the meshsize), by solving there equations which are derived from the PDE and from the fine-grid residuals; the solution of these coarse-grid equations is obtained by using recursively the same two processes. As a result, large scale changes are effectively calculated on correspondingly coarse grids, based on information gathered from finer grids. Such multigrid solvers yield *linear complexity*, i.e., the solution work is proportional to the number of variables in the system.

In many years of research, the range of applicability of these methods has steadily grown, to cover most major types of linear and nonlinear large systems of equations appearing in sciences and engineering. This has been accomplished by extending the concept of "smoothness" in various ways, finally to stand generally for any poorly locally determined solution component, and by correspondingly diversifying the types of coarse representations, to include for instance grid-free solvers (algebraic multigrid [7], [8], [9], [16]), non-deterministic problems ([10], [20], [21], [11], [12]) and multiple coarse-level representations for wave equations [5].

It has been shown (see survey [29]) that the inter-scale interactions can indeed eliminate all kinds of scale-associated difficulties, such as: slow convergence (in minimization processes, PDE solvers, etc.); critical slowing down (in statistical physics); ill-posedness (e.g., of inverse problems); conflicts between small-scale and large-scale representations (e.g., in wave problems, bridging the gap between wave equations and geometrical optics); numerousness of long-range interactions (in many body problems or integral equations); the need to produce many fine-level solutions (e.g., in optimal control, design and data assimilation problems), or a multitude of eigenfunctions (e.g., in calculating electronic structures) or very many fine-level independent samples (in statistical physics); etc. Since the local processing (relaxation, etc.) in each scale can be done in parallel at all parts of the domain, the multiscale algorithms, based on such processing, proved ideal for implementation on massively parallel computers.

To obtain even further generality, there emerge however two basic reasons to go much beyond these multigrid methods. First, they cannot perform well for *highly nonlinear cases*, where configurations cannot be decomposed into weakly-interacting local and non-local parts. Second, for many systems, even attaining linear complexity is not good enough, since the number of variables is huge. Such systems on the other hand are typically *highly repetitive*, in the sense that the same small set of governing equations (or Hamiltonian terms) keep repeating itself throughout the physical domain. This opens the way to the possibility of having, at the coarse level too, a small set of governing equations that are valid everywhere, and that can be derived from fine-level processing conducted only in some small representative "windows" (see below).

These two basic reasons point in fact in the same direction. Instead of relaxing the given system of equations so as to obtain a smooth error that can be approximated on a coarse level, one should use coarse level variables that are little sensitive to relaxation (e.g., representing chosen *averages*, rather than a subset of individual fine-level values) and that represent the *full* solution rather than the correction to any given current approximation. Such coarse variables can be chosen (as described below) so that the coarse-level equations can be derived just by local processing. We use the term *"upscaling"* for this type of direct (full-solution) transition from a fine level to a coarser one. Such a transition is valid even in those highly nonlinear cases where all scales interact with each other so strongly that correction-based multi-leveling is inapplicable.

In fact, upscaling, under the name "renormalization", was first introduced into exactly those systems where all scales interact most strongly: systems of statistical mechanics at the critical temperature of phase transition. The *renormalization group* (RG) method (see, e.g., [17], [18], [19], [65], [66]) was developed contemporaneously with, but independently of the multigrid method, its chief purpose having been to investigate the behavior of such critical systems at the limit of very large scales. The RG method has thus mainly focused on analyzing, theoretically and computationally, the fixed point of the group of successive renormalization steps, and various universal asymptotic power laws associated with it. Little has been done to upscale systems without a fixed point, which is the prevalent situation in many practical problems. This is related to the fact that the RG computational efficiency remained very limited, due to the lack of a systematic coarse-to-fine transitions, which is needed either for accelerating simulations at all levels (as in multigrid solvers) and/or for confining them to small representative windows (as described below).

Systematic Upscaling. Building on the complementary advantages of multigrid and RG described above, *Systematic Upscaling (SU)* is a methodical derivation of equations (or statistical rules) that govern a given physical system at increasingly larger scales, starting at a microscopic scale where first-principle laws are known, and advancing, scale after scale, to obtain suitable variables and operational rules for processing the system at much larger scales. Unlike classical RG, the SU algorithms include repeated coarse-to-fine transitions, which are essential for (1) testing the adequacy of the set of coarse-level variables (thus providing a general tool for constructing that set); (2) accelerating the finer-level simulations; and, most importantly (3) confining those simulations to small representative subdomains (called *windows*) within the coarser-level simulations. (SU was described briefly in [68] and at length in [69].)

Difference from ad-hoc multiscale modelling. Upscaling should not be confused with various methods of *multiscale modelling* (also called "multiscale simulation") being developed in several fields (e.g., materials science). Those methods study a physical system by employing several different ad-hoc models, each describing a very different scale of the system. Their basic approach is the *fine-to-coarse parameter passing*, in which data obtained from simulating a finer scale model, often coupled with experimental observations, are used to determine certain parameters of a larger scale model, regarding the latter as a coarse graining of the former. The most basic feature missing in this approach (as in RG) is the accurate transition from coarse levels back to finer ones, and the use of this transition as a systematic vehicle for choosing an adequate set of coarse variables and for iterating between the levels to accelerate fine-level calculations or restrict them to small windows. Successful as multiscale modelling methods are in many cases, they lack generality, are often inapplicable (e.g., when the scales are not widely separated) or inaccurate (based on questionable large-scale models), and involve much slowdown due to large scale gaps. SU, by contrast, inherits from multigrid and RG general and methodical procedures to construct and iteratively employ all intermediate scales and thus attain slowness-free efficiency and fully-controlled coarse-level accuracy.

Other numerical upscaling methods. The research on numerical upscaling, i.e., precise numerical derivation of coarse-level (e.g., macroscopic) equations from fine-level laws, has had of course a long history. In particular it has been active in computational mechanics for at least 30 years. Many systems of "homogenization", i.e., rigorous derivation of macroscopic (continuum or discretized-continuum) equations from microscopic (either continuum, or discrete continuum, or atomistic) laws have been advanced, first in the engineering literature and then in more rigorous mathematical analyses.

Most widely developed are methods of *asymptotic expansion*, in particular multiple-scale asymptotic expansion for *periodic* heterogeneous structures; see, e.g., [57], [58], [59]. For elastodynamics of composite materials, high-order methods were developed, introducing effects such as polarization, dispersion and attenuation of a *single-frequency* stress waves [56], [60]. Such expansions for *initial/boundary-value* problems in periodic media were developed by Jacob Fish and his group ([62], [63], [64]), who then extended the theory to discrete-state (atomistic) fine-scale models [42], and then also to finite temperatures [43].

Asymptotic expansions, not assuming fine-scale periodicity but relying on vast scale separations, were developed for singularly perturbed systems of differential equations based on rigorously derived averaging principles, both for deterministic and stochastic problems ([46], [47], [48], [49] for example; see comprehensive survey in [67]). The usefulness of these theories was limited since "it is often impossible, or impractical, to obtain the reduced equations in closed form... This has motivated the development of algorithms such as projective and coarse projective integration ([50], [51], [52]) within the so-called equation-free framework ([53], [54]). In this framework, short bursts of appropriately initialized fine-scale simulations are used to estimate on demand the numerical quantities required to perform scientific computing tasks with coarse-grained models (time derivatives, the action of (slow) Jacobians, and, for the case of *stochastic* coarse-grained models, the local effective noise drift and diffusivity [55])." (Quoted from [45].)

The developed method is in fact not entirely free of any equation. It actually tends to be rather similar to the *Heterogeneous Multiscale Method* (HMM) [44], which numerically solves continuum equations *of a known form* but with certain unknown local data (e.g., the stress - in a fluid dynamics macroscopic model), finding the latter by performing atomistic simulations on tiny subdomains, with constraints (e.g., the local average gradient) supplied by the macroscopic model.

The "equation-free" and HMM methods require re-derivation of the macroscopic quantities by running the microscopic model at each spatial gridpoint within each time step (with grid and time step resolution fine enough for accurate macroscopic simulation). So they are not exactly "upscaling" methods in the strict sense defined above. More important, such schemes would usually require *double* scale separation as well as extensive work at each fine-level simulation. Indeed, if the fine (temporal or spatial) step is h_f, and the scale of desired coarse (macroscopic) resolution is h_c, there usually exists also at least one intermediate scale $h_i \gg h_f$ at which the equations homogenize (and equilibrate, in the stochastic case). Now h_i must of course be much smaller than h_c for this multiscale processing to pay off. Moreover, one would typically need at least $O(h_i^2/h_f^2)$ time steps (or iterations, in static or equilibrium problems) at each fine-level simulation to reliably calculate the coarse coefficients in the HMM methods, and sometimes even up to $O(h_c^2/h_f^2)$ such steps to capture the coarse-level derivatives needed in the equation-free framework (unless one uses multigrid-like accelerations, which means introducing more computational levels between h_f and h_i, or sometimes even between h_f and h_c).

A general approach for deriving coarse-level from fine-level equations, which does not require scale separation, is based on interpolation. This was originally introduced in the multigrid framework, including in particular in its nonlinear (FAS) version [2]. The approach has been applied to the case of atomistic (not PDE) fine level, for example in the framework of the *quasi-continuum (QC)* method, used in particular to derive coarse equations at subdomains where the continuum equations break down (e.g., at the tip of a crack).

It is worth noting that early versions of the QC method had previously appeared in the multigrid literature: the method of local refinements of the PDE discretization in which each finer level locally corrects the equations of the next coarser level is part of the FAS approach (see, e.g., Secs. 7–9 in [2]), and the particular situation in which the finest level is atomistic, is described for example in Sec. 1.1 of [11].

As shown, for example by Fish and Chen [61], the efficiency of interpolation-based methods depends critically on the interpolation scheme. This issue has in fact a long history in the multigrid literature, starting with [37], [41], [7], [8], [9], [16], and all later papers on Algebraic Multigrid (AMG). Still more general and accurate methods for deriving interpolations, based on relaxed vectors and Bootstrap AMG (see Sec. 17.2 in [29]), or on adaptive smooth aggregation [40], were developed in recent years.

The interpolation-based multiscale methods were extended to simple finite-temperature models in statistical mechanics ([10], [12], [38], [39], reviewed in Sec. 13.1 of [29]), showing the feasibility of calculating *thermodynamic limits* (observable averages at system sizes tending to infinity) at "*statistically optimal" efficiency* (achieving accuracy ε in $O(\varepsilon^{-2})$ random number generations!). However, attaining similar performance for the highly nonlinear models of statistical physics (involving inseparable interaction scales) by interpolation-based methods proved impossible (as explained in Sec. 13.1 of [29]), which gave rise to the Renormalization Multigrid (RMG) method ([6] or Sec. 13.2 of [29]), the forerunner of the SU methods described here.

In the SU method, no scale separation is assumed; in fact, small ratio between successive scales can often be essential: it ensures slowdown-free computations that at each scale can be confined to certain representative windows, each containing a moderate, bounded number of variables.

Notice that SU does not depend on having a coarse level of continuum type (see examples below). Moreover, unlike other approaches, SU is being developed also for *non-local* interactions (see Sec. 6.1 below). Note as well that the SU coarsening can also be used for vast *acceleration* of fine level calculations, sometimes even by just one modest coarsening step (e.g., with just 1:3 coarsening ratio: see Sec. 2.1 below).

A basic feature of SU is that coarsening is derived *once for all*. That is, the life of each window is limited: once it has accumulated enough statistics (translated into coarser level equations) its task is over. The total size-times-duration of all the windows serving a given system at a given level depends only on the number of *different* local situations that can arise, not on the overall size and duration of the system. This implies enormous potential savings not only in computer resources, but also in human effort. Because, coarsened equations developed by one can be used by others, possibly without ever returning to the fine level. One may, for example, just develop the first level of coarsening a macromolecule (or even just part of it), thereby already providing it with much faster simulations; another investigator can then build on this the next coarser level (or add the coarsening of another part), and so on. *The upscaling of an important physical system can thus become a gradual and collaborative systematic endeavor.*

The development of SU methods should provide necessary tools for surmounting extreme computational bottlenecks in many areas of science and engineering, such as: statistical mechanics, especially at phase transition; elementary-particle physics; electronic structure of molecular systems (for deriving the inter-atomic force fields); molecular dynamics of fluids, condensed matter, nano structures and macromolecules, including proteins and nucleic acids; materials science; turbulent fluid dynamics; and global optimization of systems with multiple-scale energy barriers.

Plan of this article. Section 2 outlines the basic procedures of the SU methodology, in terms of examples from partial differential equations, molecular and macromolecular dynamics, and statistical dynamics. It also reports some preliminary results. Section 3 gives more precise details of the derivation of coarse equations, using for illustration one particular example (polymer in equilibrium). Section 4 describes the creation of windows and the algorithmic flow between windows at various levels. Sec. 5 mentions some special situations, such as boundaries. Sec. 6 is a brief survey of possible extensions of the SU methodology to deal with long-range interactions, time-dependent systems, complex fluids, low temperatures and global optimization in systems with multiscale attraction basins, and transition from quantum mechanics to molecular dynamics.

2 Systematic Upscaling (SU): An outline

Local equations and interactions. Computationally we will always deal with a discrete system, whose n variables (or unknowns) $u_1, u_2, \dots, u_n$ will typically be either the discrete values of functions (grid values, or finite elements, etc.), or the locations of particles. An equation in the d-dimensional physical space (usually $1 \le d \le 4$) is called *local* if it involves only $O(1)$ neighboring unknowns. A discretized partial differential equation, for example, is a system of local equations. Similarly, an *"interaction"*, i.e., an additive term in an energy functional or Hamiltonian H, is called *local interaction* if it involves only $O(1)$ neighboring variables. In equilibrium calculations we will usually assume H to already include the $(k_B T)^{-1}$ factor, so that the probability density $P(u)$ of each configuration $u = (u_1, u_2, \dots, u_n)$ is proportional to $\exp(-H(u))$.

For simplicity of discussion we describe SU first for systems which are *stationary or at equilibrium*, and such that their equations or interactions are *local.* We will however point out in Sec. 6 extensions to long-range equations or interactions, and to dynamic and non-equilibrium systems.

Coarsening. Similar to multigrid, SU is based on two processes: The usual *local processing* (relaxation in deterministic problems, Monte Carlo (MC) in stochastic ones) and repeated *coarsening*, creating increasingly coarser descriptions of the same physical system, with inter-scale interactions to be described below. At each coarsening stage, one constructs from a current level of description (the fine level) a coarser level, employing the following general principles.

To each fine-level configuration $u = (u_1, \ldots, u_n)$ one defines (using the criterion below) a unique coarse-level configuration $Cu = u^c = (u^c_1, \ldots, u^c_m)$, which is a vector with a reduced number of variables; typically the scale ratio is $1.5 < n/m < 10$ (except of course for problems with widely separated scales, where the equations homogenize, or a renormalization fixed-point emerges, already at some intermediate scale. Introducing additional computational scales is then needed only between the finest scale and that intermediate scale. How to choose this fine-to-coarse transition C is of course a central question, which we discuss in detail below, after giving several examples and defining first the coarse-to-fine transition.

Examples of such fine-to-coarse transformations C:

(i) For discretized continuous (e.g., PDE) problems — each coarse variable is usually an average of several neighboring fine variables. (See Sec. 5 for an example where averaging only works only up to a certain scale, above which more elaborate variables should be added.)

(ii) For a simple polymer, which consists of a chain of n atoms at the three-dimensional locations $(u_1, u_2, \ldots, u_n)$ — each coarse-level "atom" location u^c_j is at the average location of q, say, consecutive real atoms:

$$u^c_j = \frac{1}{q}(u_{qj-q+1} + u_{qj-q+2} + \cdots + u_{qj}) \ , \ (j = 1, \ldots, m; \ m = \frac{n}{q}) \ . \tag{1}$$

(iii) For a simple atomistic fluid, described by the positions u in space of its n molecules — the coarse level variables are defined at the points of a lattice placed over the flow domain, with each variable u^c_j summarizing a property of the set of molecules around that lattice point (e.g., their total mass, or density, total dipole moment, etc.). Or each u^c_j may be a *vector* which summarizes *several* such properties. At lower temperatures, as the fluid starts to solidify, additional types of coarse variables must enter, accounting for larger-scale order parameters (see Sec. 6.6).

(iv) For a lattice of Ising spins — each coarse variable is again an Ising spin, standing for the *sign* of the sum of a block of fine-level spins.

Generalized interpolation. To any given coarse configuration $U = (U_1, \ldots, U_m)$, there are generally many fine-level configurations u which are *compatible* with U (i.e., such that $Cu = U$). The interpolation (transition from U to one specific fine configuration u) is created by *compatible Monte Carlo* (CMC) (or compatible relaxation, in the deterministic case), i.e., by the local processing, restricted to configurations compatible with U. The interpolation is completed once the CMC has practically reached its equilibrium (or the compatible relaxation has converged).

For instance, in the case of polymer (Example (ii) above), if the coarse variables are defined by (1), each step in a CMC would offer a simultaneous change of two consecutive atomic positions, u_k and u_{k+1}, such that $u_k + u_{k+1}$ is kept unchanged ($qj - q + 1 \leq k \leq qj - 1$; $1 \leq j \leq m$).

The CMC interpolation was first introduced for the case of lattice Ising spins, establishing its fast equilibration [6].

The general coarsening criterion. *The fine-to-coarse transformation C is said to be* adequate *if (and to the extent that) a compatible Monte Carlo that equilibrates fast (or a compatible relaxation that converges fast) is available.*

In our polymer example, for instance, if the polymer force field includes the usual bond length, bond angle and dihedral (torsion) interactions, it turns out that the coarsening (1) is adequate for $q = 2$ or 3, but a larger q yields much slower CMC equilibration. (This result makes sense: at coarsening ratio $q \leq 3$, the fixed coarse values u_j^c implicitly nearly fix the dihedral values, which determine the large-scale polymer configuration.)

In the framework of Algebraic Multigrid (AMG), the criterion of fast convergence of compatible relaxation as a tool for choosing the coarse variables (introduced in [36]) is already used by several groups. (Particularly useful is the fact that the rate of convergence of (suitably arranged) compatible relaxation is an accurate predictor of the obtainable multigrid convergence rate; so accurate in fact that it can be used to design and debug the AMG solver.)

A major problem in coarsening any system is to find a suitable set of coarse variables. The above criterion gives a general and very effective tool for developing such a set. Importantly, a coarsening that satisfies that criterion practically implies local dependence of every fine variable on neighboring coarse variables, and hence the theoretical possibility to construct, just by local processing, a set of *"equations"* (in the form of numerical tables) that will govern simulations at the coarse level. How to go about this construction in practice is discussed in Sec. 3 below.

Moreover, in highly repetitive systems (defined above), this local construction of the coarse equations need not be done everywhere: the coarse-level equations can iteratively be derived by comparing coarse-level with fine level simulations, where the latter are performed only in some relatively small *windows* (subdomains, on the boundaries of which the fine level is kept compatible with the coarse level. See Sec. 4 below for details.)

Thus, the fine level simulations supply (or correct) the equations (or Hamiltonian) of the next coarser level. On the other hand, the coarse level selects the windows where these fine-level simulations should take place. A window is opened over a region where the coarse level has detected local relations in a range for which fine level computations in previous windows could not supply accurate equations (see more in Sec. 4). Iterating back and forth between windows at all the levels quickly settles into a self-consistent multilevel equilibrium and compatibility; as in multigrid, if the coarsening ratio n/m is not large, no slowdown should occur. More important, at each level the computations need extend only over a collection of small windows, whose number depends on the diversity of local situations, not on the size of the entire system.

2.1 Experimental results from the four simple examples mentioned above have already revealed the very high potential of the SU approach. For instance, in Example (ii), in which conventional simulations run into extreme slowdowns, even the single coarsening level (1), with $q = 3$, already accelerates the simulation by at least two orders of magnitude, while accurately reproducing all the relevant statistics (using the coarse Hamiltonian described in Sec. 3 below. This result was obtained in computations devised by Dr. Dov Bai). The reason for this acceleration is that this first coarsening already effectively averages over the attraction basins caused by the local minima of the fine-level dihedral interactions.

Also, as mentioned, wide experience has been accumulated confirming the effectiveness of the above general coarsening criterion in the special case of Algebraic Multigrid. The criterion has also been essential in developing coarse variables for high-Reynolds flows (see Sec. 5 below).

3 Derivation of coarse equations

Basic hypothesis: localness of coarsening. The *solution* to a system each of whose equations is local cannot be determined locally: it depends on *all* equations, near and far. However, *what can essentially be determined just from local information are coarser-level equations (or interactions).* More precisely: provided the coarse set of variables is adequate (satisfying the above general coarsening criterion), a coarse system of local equations (interactions) equivalent to the fine-level system (in the sense that a coarse solution/equilibrium would yield the fine solution/equilibrium by a brief compatible local processing) is obtainable locally (i.e., by processing only a fine-level neighborhood comparable in size to the typical distance between coarse variables) with an error that decreases exponentially as a function of the total work, assuming this work is invested to suitably enlarge the number of variables involved in each coarse equation and increase the size of the fine-level neighborhood and the number of iteration involved in the local processing. This hypothesis has emerged from the long and diverse experience with both RG and multigrid solvers.

The actual derivation of the coarse equations, incorporating RG-type techniques, is based on fine-level simulations in representative regions. The simulations give a sequence of fine-level configurations u, which is readily translated into a sequence of coarse configurations $u^c = Cu$. There exist several approaches as to how and in what form to derive governing coarse-level rules from such a sequence. We briefly describe two basic approaches with which some experience has already been gained (recommending mainly the second of them).

3.1 Dependence table. In this (older) approach, the sequence of coarse configurations calculated by fine-level simulations is used to accumulate statistics of the dependence of each coarse variable (the *"pivot"*) on a certain set of neighboring coarse variables (the *"neighborhood"*). For this purpose the set of possible neighborhood configurations is partitioned into bins (described below). In fully deterministic problems, the average value of the pivot in each bin is accumulated and then tabulated. From such a table, the pivot value for any given neighborhood configuration can be interpolated, which is all one needs in order to operate (e.g., perform Gauss-Seidel relaxation) at the coarse level. In stochastic problems, additional statistics (e.g., variance) of the pivot over each neighborhood bin are tabulated, enough to enable accurate Monte Carlo simulations at the coarse level. Successful experience with simple versions of this approach, including the cases of Examples (iii) and (iv) above, are reported in [6], [24], [25] and [26].

The partition into bins should be done in terms of neighborhood functionals $S_1, S_2, \ldots$, roughly ordered in decreasing order of their influence on the pivot. For example, S_1 can be the sum of the values at the nearest neighbors to the pivot; S_2 — the sum of the next-nearest neighbors; S_3 — the variance in the set of nearest-neighbor values; etc. The space $S = (S_1, S_2, \ldots)$ is divided into (very roughly equi-probability) bins. The number of bins need not be very large; it only needs to allow sufficiently accurate interpolation to any probable value $s \in S$. The interpolation is usually done by a polynomial whose integral over each bin near s yields the statistics accumulated at that bin; the interpolation accuracy can of course deteriorate with decreasing probability of s.

The construction of the space of functionals S and its binning should be guided by physical insight. It need not be very particular: many different choices would be adequate, due to the interdependence between neighborhood functionals. Still, the construction will usually be cumbersome. Therefore, for most models, the next approach seems preferable, so we describe it in a greater detail.

3.2 Coarse Hamiltonian. In this approach the sequence of coarse configurations is used to calculate averages of quite many coarse-level observables $O_1, O_2, \ldots, O_\mu$. Denote the average of O_i by $\langle O_i \rangle_f$, subscript f indicating that this average of the coarse observable has been calculated in the *fine*-level simulations. The coarse level itself is intended to be governed by a (yet to be calculated) Hamiltonian-like functional $H^c(u^c)$, that is, the probability of a coarse configuration u^c will be proportional to $\exp(-H^c(u^c))$. For any given approximate H^c, one can run simulations at the coarse level during which $\langle O_i \rangle_c$, the average of O_i according to H^c, can be calculated. Our first aim is to construct H^c such that $\langle O_i \rangle_c = \langle O_i \rangle_f$, $(i = 1, \ldots, \mu)$.

For this purpose H^c is written in the general form

$$H^c(u^c) = \sum_{k=1}^{K} a_k H_k(u^c) \ , \tag{2}$$

where each H_k is a known functional of u^c (see examples below) and $\{a_k\}$ is a set of coefficients that need to be found. A crude approximation to H^c, possibly with a reduced number of terms (reduced K), can inexpensively be obtained from small-scale fine-level calculations assuming independence of various quantities (as in the example below). The approximation is then improved in few Newton-like iterations, during which additional terms may be added to H^c (increasing K). Specifically, in each iteration H^c is changed by adding to it $\delta H^c = \sum_k \delta a_k H_k$. Using the first-order relation (used already by K.G. Wilson and R.H. Swendsen; cf., e.g., [27])

$$\delta\langle O \rangle = \langle O \rangle \langle \delta H^c \rangle - \langle O \cdot \delta H^c \rangle \ , \tag{3}$$

one gets a system of μ equations

$$\sum_{k}^{K} (\langle O_i \rangle \langle H_k \rangle - \langle O_i H_k \rangle) \delta a_k = \langle O_i \rangle_f - \langle O_i \rangle_c \ , \ (i = 1, \ldots, \mu) \tag{4}$$

where $\langle O_i \rangle$, $\langle H_k \rangle$ and $\langle O_i H_k \rangle$ are averages calculated during the fine-level simulations. (Corresponding averages are also calculated in the coarse-level simulation, for a purpose explained below.) the corrections $\{\delta a_k\}$ at each iteration are calculated to satisfy (4) best, in a least-square sense. We choose for this purpose enough observables: $\mu > K$. Usually, the set of observables $\{O_i\}_i$ will include $H_1, H_2, \ldots, H_K$ and some others. Then H^c is replaced by $H^c + \delta H^c$ and if this change is not small enough, another iteration is performed with the new H^c.

These Newton-like iterations converge fast. Having converged, the relation $\langle O_i \rangle_c = \langle O_i \rangle_f$ is satisfied for $i = 1, \ldots, \mu$, but possibly not for observables which were not included in the process. To check the accuracy of the calculated H^c, we therefore calculate the discrepancy $\langle O \rangle_f - = \langle O \rangle_c$ for many additional observables, in particular for all the "second moment", or product observables $O_i H_k$, $(i = 1, \ldots, \mu;\ k = 1, \ldots, K)$, whose averages, needed in (4), have anyway been calculated. An observable O for which the discrepancy $\langle O \rangle_f - \langle O \rangle_c$ remains particularly large can be added to the list of Hamiltonian terms (increasing K) to facilitate decrease of this discrepancy in the next iterations. Due to the inter-dependence of observables, obtaining in this way accurate reproduction of the observables that have had the largest discrepancies will usually cause the discrepancies to sharply decrease in other observables as well, including still-higher-moment observables. The iterations continue as long as all such observables exhibit satisfactorily small discrepancies.

Example. A test of the coarse Hamiltonian approach has been carried out with the polymer case (Example (ii) in Sec. 2), employing a united-atom model of polymethylene taken from [28], using the coarsening (1) with $q = 3$. (A detailed description, but with $q = 4$, has appeared in [22].) The given fine-level Hamiltonian in this case is the sum of bond lengths, bond angles, bond dihedrals (torsions) and Lennard-Jones interactions. Similarly, the first approximation to H^c is chosen in the form

$$H^c(u^c) = \sum_i F_1(|r_i - r_{r+1}|) + \sum_i F_2(\theta_i) + \sum_i F_3(\tau_i) + \sum_{|i-j|>2} F_4(|r_i - r_j|) \ , \tag{5}$$

where $r_i = u_i^c$ is the location of the i-th coarse "atom", $|r_i - r_{i+1}|$ is the distance between two successive coarse "atoms", θ_i is the angle (r_{i-1}, r_i, r_{i+1}), τ_i is the torsion $(r_{i-1}, r_i, r_{i+1}, r_{i+2})$, i.e., the angle between the plane spanned by (r_{i-1}, r_i, r_{i+1}) and the one spanned by (r_i, r_{i+1}, r_{i+2}), and F_4 is a Lennard-Jones-like interaction. Each of the initially-unknown functions F_ℓ can be expanded in the form

$$F_\ell(\xi) = \sum_j a_{\ell,j} w_j(\xi) \ , \ (\ell = 1, 2, 3, 4) \tag{6}$$

with unknown coefficients $a_{\ell,j}$ and known basis function $w_j(\xi)$; e.g., local basis functions (one-dimensional finite elements). Upon collecting (over the relevant $\sum_i$ in (5)) all the Hamiltonian terms that include the same unknown $a_{\ell,j}$, the coarse Hamiltonian (5) obtains the general form (2).

A reasonable first approximation to F_4 is the given Lennard-Jones interactions, multiplied by q^2. A first approximation to F_1 (similarly: F_2, F_3) can be calculated by fine-level simulation of a rather short polymer chain (e.g., $n = 24$ and $m = 8$) from the bare distribution of the distances $\{|r_i - r_{i+1}|\}_{i=2}^{m-2}$, omitting the exceptional distances at the ends ($i = 1$ and $i = m - 1$), for which a different function F_1 may separately be calculated. These first approximations ignore all correlations between the internal coordinates ($\{|r_i - r_{i+1}|\}, \{\theta_i\}$ and $\{\tau_i\}$). The iterations described above will then automatically correct for those correlations, introducing on the way some new explicit correlation terms into H^c (increasing K by adding "second-moment" terms as described above). With the improved H^c one can then make simulations at the coarse level with much longer chains (e.g., $n = 120$) and further improve H^c by making additional iterations, now making the fine-level simulations in *windows within that longer chain* (see Sec. 4). Then H^c can be similarly used to derive Hamiltonians at still coarser levels, with simulations on still longer chains. The longer chains may produce new situations (e.g., contact points due to folding) that require some new, window-within-window calculations at all finer levels to further correct H^c. The formulation is very flexible, allowing the introduction of new Hamiltonian terms to account for new situations (see next).

4 Window Developments

Usually the computation starts in a rather small finest-level domain, with some artificial boundary conditions. For example, in the case of atomistic fluids (Example (iii) above), one can start with several hundreds (in two-dimensional problems) or several thousands (in 3D) particles, with periodic boundary conditions, where the period is chosen so that the fluid has its known average density. Similarly, in the case of discretized PDE problems (e.g., a fluid satisfying discretized Navier-Stokes equations), one can start with a grid of several hundreds (in 2D) to several thousands (in 3D) gridpoints, with periodic boundary conditions. In the polymer case, one can start with a short chain, as mentioned above.

¿From the computations in those small domains, one derives the first approximation to the first-coarse-level equations (or Hamiltonian), as in the example above. Being coarser, the new level then allows inexpensive calculations in a much larger domain. Note, however, that these computations in a larger domain are likely to encounter situations not accounted for by the initial small-domain fine-level simulations. For instance, the density fluctuations in the large domain can be much larger than in the small domain, creating regions with densities outside the ranges that have been simulated (or that have accumulated enough statistics) at the fine level. Similarly, in the polymer example, the short initial chain may well have lacked folding situations typical to longer chains. To obtain more accurate coarse equations in such new situations, more fine-level small-domain calculations should be done.

Some of the needed additional calculations could in fact have been done in advance. Namely, *several* small-domain calculations could initially be launched, independent of each other, to simulate for example several different average densities, or several different mixtures of atomic species or chain elements. However, it is difficult to anticipate in advance all the different local situations that will arise. So, generally, much of the needed additional fine level simulations will be carried out "upon demand", in regions where the coarse level encounters new situations, and they will be carried out as fine level *windows* within the coarse level simulation. This has also the advantage of giving these simulations more realistic boundary conditions (for fluids), or more realistic folding situations (for polymers), etc.

A window is created by replacing a certain coarse subdomain (where new situations have arisen) by a fine-level patch (including again several hundreds to several thousands degrees of freedom). On the boundary of that patch the fine-level simulations are kept compatible with the coarse simulations using the generalized interpolation described above. As in Sec. 3 above, the fine simulation will give rise to (or correct) the coarse equations characteristic to the refined subdomain.

The coarse-level equations derived in the window can of course be used outside that window's subdomain, wherever similar conditions exist. Also, the window may be shut off (returning to pure coarse level simulations at that subdomain) as soon as it has accumulated enough statistics to make the derived course equations as accurate as desired.

The process can of course be recursive: The coarse-level simulations can be used to construct equations for a still coarser level, which will be simulated in a still larger domain, possibly creating, upon encountering additional local conditions, new windows of the first coarse level, with new fine-level windows inside them. And so on.

5 Some Special Situations

Various special situations require special or modified coarse-level equations (or additional terms in H^c). In the case of PDEs, special coarse equations would usually be needed near boundaries, with different equations near different types of fine-level boundary conditions.

In some cases, new types of variables need be introduced to satisfy the above general coarsening criterion. For example, in Navier-Stokes simulations of incompressible two-dimensional fluids, at fine enough scale suitable coarsening can be defined in terms of averages of the vorticity function ω. But this coarsening is no longer adequate for large time steps and/or large spatial meshsizes that do not sufficiently resolve the rotation of the flow in strong vortexes. It has been shown that suitable coarsening (satisfying the above coarsening criterion) at such scales can be constructed by decomposing the vorticity function into the sum of idealized vortexes (radial local functions that satisfy the steady-state Euler equations) and the averages of a remaining low-vorticity part (work in progress, in collaboration with Jim McWilliams, Boris Diskin and

Young Ju Lee).

In the polymer case described above, each of the coarse-level functions F_l should have a separate special expansion (6) for the internal-coordinates near the ends of the chain. Hence these coordinates will have different sets of unknowns $a_l j$, which thus cannot be collected together with the terms arising away from the ends. Also, at large scales, special situations arise when the polymer folds upon itself, bringing into proximity atoms that are not neighbors along the chain. Special terms should then be added to H^c that depend on the distances between such atoms and on local angles created by them.

6 Extensions in Brief

Various important extensions of the upscaling techniques, to diverse physical situations, can be developed. The following principal directions have already been preliminarily considered.

6.1 Long range interactions (e.g., between electrostatic charges) can each be decomposed into the sum of a smooth interaction and a local one ("smooth" and "local" being meant on the scale of the next coarse level; all familiar physical interactions can be decomposed this way; see [23] and examples in [30], [33] and [32]). To any desired accuracy, the smooth part can directly be represented at the coarse level, e.g., by aggregated charges and dipoles moving with the coarse level "atoms" (in Example (ii) above) or by (high order) adjoint interpolation of charges to the coarse-level lattice (in Example (iii)). The local part is essentially transferred, together with all other local interactions, using the fine/coarse iterations described above. Effectively, the amount of work invested per charge involves only calculating its local interactions, and, even more importantly, only charges within selected windows need be treated. It can be shown that this is possible due to the smoothness of the non-local interactions, which makes them minimally sensitive to the local MC moves, especially when the latter are explicitly designed to conserve certain moments of the charge distribution.

6.2 Dynamical systems. Generally, for time-dependent systems, the equilibrium coarsening criterion of Sec. 3.1 is replaced by the requirement that the fine-level configuration (if its evolution is stable) or its ensemble statistics (in the case of instability) at any given time can fully be recovered from the coarse configurations at a small number of previous time steps. Dependence tables (e.g., in kinetic Monte Carlo computations) have been derived in the form of flux dependence on both current-time and previous-time neighboring coarse variables [34]. A general computational criterion has been formulated for the size of the time steps to increase with the spatial coarsening level, so as to maintain full efficiency. A Hamiltonian-like functional that governs every time step can also be developed analogously to the one described in Sec. 3.

For Hamiltonian systems (e.g., corresponding to Examples (ii) and (iii) in Sec. 3.1), the multiscale structure allows a natural combination of temperature-accurate statistical simulation at small scales with time-accurate dynamics at large scales. Assuming that after any given time interval the fine-scale degrees of freedom settle into a local equilibrium slave to the coarse-level averages (where that level increases with the size of the time interval), the general equilibrium criterion for choosing the coarse variables (see Sec. 2) still directly applies. Large time steps, based on implicit discretization of Newton law, can then be made, using a multigrid-like solver where the relaxation at fine levels is replaced by CMC (cf. Sec. 14.8 in [29]). This approach yields two benefits in performing very large time steps: firstly, it allows much easier handling of local minima. Secondly, it avoids the killing of highly-oscillatory modes (those vibrations that are not resolved by the time step), which would occur if the implicit equations of a large time step were imposed at all scales. Instead, these modes assume stochastic amplitudes according to their equilibrium probability distribution. The desired temperature is introduced very directly in this way, avoiding the need for fabricating Langevin stochastic forces.

Another possible approach, for fluids, is to first develop at equilibrium a coarse-level Hamiltonian $H^c(u^c)$ such that the relation $Probability(u^c) \sim \exp(-H^c(u^c)/k_BT)$ will simultaneously hold for a full range of temperatures T. (This can be achieved by adding several moments $(H^c)^m$ to the list of observables (O_i) used in (4), and constructing a *joint Hamiltonian* (see below) for different temperatures in the range of interest). Then use H^c in Newtonian dynamics at the coarse level, where effective coarse-level masses (and their possible dependence on the coarse coordinates) are determined by comparing (in windows of fine-level dynamic simulations) coarse-level accelerations with the gradient of H^c.

Still another approach for fluids, is a Boltzmann-type upscaling in the 6D space of positions and velocities. Starting simulations at the individual-particle level, increasingly coarser spatial levels will describe velocity distributions at progressively higher resolutions.

6.3 Stochastic coarsening. Our studies (e.g., [34]) have shown that averaging upon coarsening should often better be stochastic. The added stochasticity is important to create smoother coarse dynamics, hence simpler dependence table or easier H^c, especially for a fine level with discrete-state (e.g., integer-valued) variables or highly oscillating Hamiltonian. One general way is to modify a deterministic averaging (or anterpolation — the adjoint of interpolation) by adding to each coarse variable a small increment, where the field of increments is in equilibrium governed by a Hamiltonian-like functional H_p. A corresponding CMC process has been developed, and the general coarsening criterion then effectively checks that H_p has been properly designed, i.e., so that it prohibits increment fields that correspond to long-range moves.

6.4 Joint H^c. The same coarse functional H^c should sometimes simultaneously satisfy $\{< O_i >_c = < O_i >_f\}_i$ for several different MC situations, such as: (a) under different external fields; (b) at different temperatures (cf. Sec. 6.2); (c) in different energy basins (cf. Sec. 6.6). Generally, this can be achieved by adding in (2) terms H_k that are particularly sensitive to the differences between the different simulated situations.

6.5 Complex fluids. More elaborate coarse Hamiltonians are needed for fluids with more complex molecules of one or several species, such as water with methanol, or glycerol, etc. A gradual construction can then be planned, starting for example with H^c constructed for atomistic equilibrium in a periodic domain containing only two molecules. Adding then to the simulation one molecule at a time, the coefficients of H^c are updated by (4), with additional terms H_k that correspond at each iteration to correlation observables that are still ill approximated.

6.6 Low temperatures (example). At high temperatures, the coarse variables for a simple fluid in equilibrium are gridpoint values, each standing for some local averaging of $m(\mathbf{x})$, the masses m of particles at positions $\mathbf{x} = (x_1, x_2, x_3)$ near the gridpoint. At lower temperatures, as the fluid starts to crystallize, roughly with periods $u^{(\ell)} = (u_1^{(\ell)}, u_2^{(\ell)}, u_3^{(\ell)})$, $(\ell = 1, 2, 3)$, say, three new coarse-level fields should enter, standing for local averaging of $\exp(2\pi i w^{(\ell)}\mathbf{x}) * m(\mathbf{x})$, $(\ell = 1, 2, 3)$, where $\mathbf{w}^{(\ell)} \cdot \mathbf{u}^{(\ell)} \simeq \delta_{k\ell}$. If the crystal is perfect and $\mathbf{w}^{(\ell)}$ are exactly known, these coarse variables will be constants. When $w^{(\ell)}$ are only approximate, these variables will oscillate smoothly. Similar averaging at the next coarser levels will then describe these oscillations, effectively correcting the erroneous $\mathbf{w}^{(\ell)}$. If the crystal is not perfect, meaningful averaging of this type will persist only up to a certain scale; usually, the lower the temperature the larger that scale (See Sec. 14.7.3 in [29]).

Also upon lowering the temperature, energy barriers emerge at increasingly larger scales. By insisting on constructing, level after level, *joint* H^c (see Sec. 6.4), statistically correct transitions between different energy basins can efficiently be simulated.

6.7 Multiscale annealing. Thus, as a system is gradually cooled, increasingly larger-scale degrees of freedom are identified. This identification of increasingly larger collective moves makes such a computation extremely more effective than simple simulated annealing [35] for *minimizing* the energy (the limit $T \to 0$), especially in the physically common situation of multiscale nested attraction basins. The multiscale annealing may provide efficient solvers to very difficult global optimization problems. (See more in Sec. 18.2 in [29], and [31].)

6.8 Coarse-level computability of fine observables. Often, an observable of interest is not directly expressed in terms of the coarse-level variables. We have developed a general procedure (similar to Sec. 3.2) for computing a functional dependence of a quantity of interest upon the coarse variables, based on suitable statistics accumulated during the fine-level simulations.

6.9 Determinism and stochasticity. The discussion above is written mainly in terms of stochastic systems, but can be extended to deterministic ones. Moreover, a stochastic system at the fine level often yields a deterministic system at large enough scales. The opposite exists too: A deterministic fluid flow at the viscous scale acquires stochastic features at the large scales reigned by turbulence. Similarly, a particle system governed by Newtonian dynamics can give rise to various stochastic developments at different scales. The coarsening approaches described above can accommodate such transitions.

6.10 Upscaling from quantum mechanics to molecular dynamics. The electronic distribution of a molecular system with N valence electrons can be approximately computed by solving the Kohn-Sham equations. (This involves the calculation of N eigenfunctions of a Schroedinger operator, whose potential depends on the eigenfunctions. Very efficient multigrid is developed for this task; see Sec. 9 in [29].) It is estimated that such calculations can solve sufficiently large systems to enable upscaling to molecular dynamics (MD) or molecular equilibrium (ME), ie., derivation of the MD or ME force fields. The coarse-level variables in this upscaling are quite obviously the nuclear positions. The force field tables for ME calculations, for example, can iteratively be derived by comparing the ME simulations with Kohn-Sham solutions, where the latter need be computed only in relatively small windows. This is similar to the derivation of the *coarse-level* macromolecular force field described in Sec. 3.2.

An even more intriguing possibility is the attempt to use the SU approach to derive the molecular force fields directly from the fundamental, high-dimensional equations of quantum mechanics, using Feynman's path integrals, as sketched in [70].

References

[1] Brandt, A., Multi-level adaptive technique (MLAT) for fast numerical solutions to boundary value problems, in *Proc. 3^{rd} Int. Conf. on Numerical Methods in Fluid Mechanics*

(Cabannes, H. and Temam, R., eds.), Lecture Notes in Physics 18, Springer-Verlag, 1973, pp. 82–89.

[2] Brandt, A., Multi-level adaptive solutions to boundary value problems, *Math. Comp.* **31** (1977) 333–390.

[3] Brandt, A., Guide to multigrid development, in *Multigrid Methods* (Hackbusch, W. and Trottenberg, U., eds.), Springer-Verlag, 1982, pp. 220–312.

[4] Brandt, A., *Multigrid Techniques: 1984 Guide, with Applications to Fluid Dynamics*, 1984, 191 pages, ISBN-3-88457-081-1; GMD-Studien Nr. 85. To appear in SIAM Classics in Applied Mathematics series, 2007.

[5] Brandt, A. and Livshits, I., Wave-ray multigrid method for standing wave equations, *Electronic Trans. Num. An.* **6** (1997), 162–181.

[6] Brandt, A. and Ron, D., Renormalization multigrid (RMG): Statistically optimal renormalization group flow and coarse-to-fine Monte Carlo acceleration, Gauss Center Report WI/GC–11, June 1999. *J. Stat. Phys.* **102** (2001) 231–257.

[7] Brandt, A., McCormick, S. and Ruge, J., Algebraic multigrid (AMG) for automatic multigrid solution with application to geodetic computations, Institute for Computational Studies, POB 1852, Fort Collins, Colorado, 1982.

[8] Brandt, A., McCormick, S. and Ruge, J., Algebraic multigrid (AMG) for sparse matrix equations, in *Sparsity and its Applications* (Evans, D.J., ed.), Cambridge University Press, Cambridge, 1984, pp. 257–284.

[9] Brandt, A., Algebraic multigrid theory: The symmetric case, in *Preliminary Proc. Int. Multigrid Conf.*, Copper Mountain, Colorado, April 6–8, 1983; *Appl. Math. Comp.* **19** (1986) 23–56.

[10] Brandt, A., Ron, D. and Amit, D.J., Multi-level approaches to discrete-state and stochastic problems, in *Multigrid Methods, II* (Hackbusch, W. and Trottenberg, U., eds.), Springer-Verlag, 1986, pp. 66–99.

[11] Brandt, A., Multigrid methods in lattice field computations, *Nucl. Phys. B* (Proc. Suppl.) **26** (1992) 137–180.

[12] Brandt, A., Galun, M. and Ron, D., Optimal multigrid algorithms for calculating thermodynamic limits, *J. Stat. Phys.* **74** (1994) 313–348.

[13] Hackbusch, W., Multigrid Methods and Applications. Springer, Berlin, 1985.

[14] Trottenberg, U, Oosterlee, C.W., and Schüller, A., *Multigrid*. Academic Press, London, 2000.

[15] Briggs, W.L., Henson, V.E., and McCormick, S.F., *A Multigrid Tutorial*, 2nd Ed., SIAM, 2000.

[16] Ruge, J. and Stüben, K., Algebraic multigrid. In *Multigrid Methods* (McCormick, S. F., ed.), SIAM, Philadelphia, 1987, pp. 73–130.

[17] Wilson, K.G., The Renormalization Group and Critical Phenomena. *Rev. Mod. Phys.*, 55 (1983) 583–600. (1982, Nobel Prize Lecture)

[18] Swendsen, R.H., Renormalization group Monte Carlo methods, *Phys. Rev. Lett.* **42** (1979) 859–862.

[19] Swendsen, R.H., Monte Carlo renormalization-group studies of the $d = 2$ Ising model, *Phys. Rev. B* **20** (1979) 2080–2087.

[20] Mack, G. and Pordt, A., Convergent perturbation expansions for Euclidean quantum field theory., *Comm. Math. Phys.* **97** (1985) 267. Also: Mack, G. in: Nonperturbative Quantum Field Theory, G. t'Hooft *et al.*, eds. Plenum Press, New York, 1988, p. 309.

[21] Goodman, J. and Sokal, A.D., Multigrid Monte Carlo methods for lattice field theories. *Phys. Rev. Lett.*, 56 (1986) 1015–1018.

[22] Bai, D. and Brandt, A., Multiscale computation of polymer models. In: Brandt, A., Bernholc, J. and Binder, K. (Eds.), *Multiscale Computational Methods in Chemistry and Physics*. NATO Science Series: Computer and System Sciences, Vol. 177, IOS Press, Amsterdam (2001), pp. 250–266.

[23] Brandt, A., Multilevel computations of integral transforms and particle interactions with oscillatory kernels, *Comp. Phys. Comm.* **65** (1991) 24–38.

[24] Brandt, A. and Iliyn, V., Multilevel approach in statistical physics of liquids. In: Brandt, A., Bernholc, J. and Binder, K. (Eds.), *Multiscale Computational Methods in Chemistry and Physics*. NATO Science Series: Computer and System Sciences, Vol. 177, IOS Press, Amsterdam (2001), pp. 187–197.

[25] Brandt, A. and Ilyin, V., Multilevel Monte Carlo methods for studying large-scale phenomena in fluids. *J. of Molecular Liquids*, **105** (2003) 253–256.

[26] Shmulyian, S., *Toward Optimal Multigrid Monte Carlo Computation in Two-Dimensional $O(N)$ Non-Linear σ-Models*, Ph.D. Thesis, Weizmann Institute of Science, 1999.

[27] Gupta, R., Open problems in Monte Carlo Renormalization Group: Application to critical phenomena, *J. Appl. Phys.* **61** (1987) 3605–3611.

[28] Paul, W., Yoon, D.Y. and Smith, G.D., An optimized united atom model for simulations of polymethylene melts, *J. Chem. Phys.* **103** (1995) 1702–1709.

[29] Brandt, A., Multiscale scientific computation: review 2001. In Barth, T.J., Chan, T.F. and Haimes, R. (eds.): *Multiscale and Multiresolution Methods: Theory and Applications*, Springer Verlag, Heidelberg, 2001, pp. 1–96. Available in www.wisdom.weizmann.ac.il/∼achi/review00.ps.

[30] Brandt, A. and Venner, C.H., Multilevel evaluation of integral transforms with asymptotically smooth kernels, Gauss Center Report WI/GC–2, The Weizmann Institute of Science, Rehovot, April 1995; *SIAM J. Sci. Comp.* **19** (1998), 468–492.

[31] Brandt, A. and Ron, D., Multigrid solvers and Multilevel Optimization Strategies. In: *Multilevel Optimization and VLSICAD*, (J. Cong and J.R. Shinnerl, eds.), Kluwer Academic Publishers, Boston, 2003, pp. 1–69.

[32] Ramirez, I.H., Multilevel Multi-Integration Algorithm for Acoustics, Ph.D. Thesis, University of Twente, Enschede, The Netherlands, 2005.

[33] Suwan, I., Multiscale Methods in Molecular Dynamics, Ph.D. Thesis, Feinberg Graduate School, Weizmann Institute of Science, Rehovot, Israel, 2005.

[34] Saad, N., Multiscale Algorithm for Time-Dependent One-Dimensional System, Feinberg Graduate School, Weizmann Institute of Science, Rehovot, Israel, 2005.

[35] Kirkpatrick, S., Gelatt Jr., C.D. and Vecci, M.P., Optimization by simulated annealing, *Science* **220** (1983), 671.

[36] Brandt, A., General highly accurate algebraic coarsening schemes. Gauss Center Report WI/GC–13, May 1999. *Electronic Trans. Num. Anal.* **10** (2000) 1–20.

[37] Alcouffe, R.E., Brandt, A., Dendy, J.E., Jr. and Painter, J.W., The multigrid methods for the diffusion equation with strongly discontinuous coefficients, *SIAM J. Sci. Stat. Comp.* **2** (1981) 430–454.

[38] Brandt, A. and Galun, M., Optimal multigrid algorithm for the massive Gaussian model and path integrals, *J. Stat. Phys.* **82** (1996) 1503–1518.

[39] Brandt, A. and Galun, M., Optimal multigrid algorithms for variable-coupling isotropic Gaussian models. *J. Stat. Phys.* **88** (1997) 637–664.

[40] M. Brezina, R. Falgout S. MacLachlan T. Manteuffel S. McCormick and J. Ruge, Adaptive smoothed aggregation (αSA) multigrid, *SIAM Review* **47** (2005) 317–346.

[41] Dendy, J.E., Jr., Black box multigrid. *J. Comp. Phys.* **48** (1982) 366–386.

[42] Chen, W. and Fish, J., A generalized space-time mathematical homogenization theory for bridging atomistic and continuum scales, *International J. Num. Meth. in Eng.* **67** (2006) 253–271.

[43] Fish, J., Chen, W. and Li, R., Generalized mathematical homogenization of atomistic media at finite temperatures in three dimensions, *Comput. Methods Appl. Mech. Engr.* (2006) doi:10.1016/j.ema.2006.08.001.

[44] E, W. and Engquist, B., The heterogeneous multi-scale methods, *Commun. Math. Sci.* **1** (2002) 87–132.

[45] Givon, D., Kevrekidis, I.G. and Kupferman, R., Strong convergence of projective integration schemes for singularly perturbed stochastic differential systems, *Comm. Math. Sci.* **4** (2006) 707–729.

[46] van Kampen, N.G., Elimination of fast variables *Phys. Rep.* **124** (1985) 69–160.

[47] Skorokhod, A.V., Asymptotics Method in the Theory of Stochastic Differential Equations, AMS, Providence, RI, 1987.

[48] Veretennikov, A. Yu., On the averaging principle for systems of stochastic differential equations, *Math. USSR Sborn.* **69** (1991) 271–284.

[49] Freidlin, M.I. and Wentzell, A.D., Random Perturbations of Dynamical Systems, Springer, New York, 1984.

[50] Gear, C.W., Kevrekidis, I.G. and Theodoropoulos, C., "Coarse" integration/bifurcation analysis via microscopic simulators: micro-Galerkin methods, *Comp. Chem. Engr.* **26** (2002) 941–963.

[51] Gear, C.W. and Kevrekidis, I.G., Projective methods for stiff differential equations: problems with gaps in their eigenvalue spectrum, *SIAM J. Sci. Comp.* **24** (2003) 1091–1106.

[52] Rico-Martinez, R., Gear, C.W. and Kevrekidis, I.G., Coarse projective KMC integration: forward/reverse initial and boundary value problems, *J. Comp. Phys.* **196** (2004) 474–489.

[53] Kevrekidis, I.G., Gear, C.W., Hyman, J.M., Kevrekidis, P.G., Runborg, O. and Theodoropoulos, K., Equation-free coarse-grained multiscale computation: enabling microscopic simulators to perform system-level tasks, *Comm. Math. Sci.* **1** (2003) 715–762.

[54] Kevrekidis, I.G., Gear, C.W. and Hummer, G., Equation-free: the computer assisted analysis of complex multiscale systems, *AIChE J.* **50** (2004) 1346–1354.

[55] Hummer, G. and Kevrekidis, I.G., Coarse molecular dynamics of a peptide fragment: free energy, kinetics and long time dynamics computations, *J. Chem. Phys.* **118** 10762–10773.

[56] Ting, T.C.T., Dynamic response of composites, *Appl. Mech. Rev.* **33** (1980) 1629–1635.

[57] Benssousan, A., Lions, J.L. and Papanicoulau, G., Asymptotic Analysis for Periodic Structures, North Holland, Amsterdam, 1978.

[58] Bakhvalov, N.S. and Panasenko, G.P., Homogenization: Averaging Processes in Periodic Media, Kluwer, Dordrecht, 1989.

[59] Sanchez-Palencia, E., Non-homogeneous Media and Vibration Theory, springer, Berlin, 1980.

[60] Boutin, C. and Auriault, J.L., Rayleigh scattering in elastic composite materials, *Int. J. Engng. Sci.* **12** (1993) 1669-1689.

[61] Fish, J. and Chen, W., Discrete-to-continuum bridging based on multigrid principles, *Computer Methods in Applied Mechanics and Engineering*, **193** (2004) 1693–1711.

[62] Fish, J., Chen, W. and Nagai, G., Nonlocal dispersive model for wave propagation in heterogeneous media: one-dimensional case, *International Journal for Numerical Methods in Engineering*, **54** (2000) 331–346.

[63] Fish, J., Chen, W. and Nagai, G., Nonlocal dispersive model for wave propagation in heterogeneous media: multidimensional case. *International Journal for Numerical Methods in Engineering* **54** (2004) 347–363.

[64] Fish, J. and Chen, W., Space-time multiscale model for wave propagation in heterogeneous media. *Comp. Meth. Appl. Mech. Engng.* **193** (2004) 4837–4856.

[65] Kadanoff, L., Statistical Physics, Statics, Dynamics and Renormalization, World Scientific, Singapore, 2002.

[66] Fisher, M.E., Renormalization group theory: Its basis and formulation in statistical physics. *Rev. Mod. Phys.*, 70(2) (1998) 653–681.

[67] Givon, D., Kupferman, R. and Stuart, A., Extracting macroscopic dynamics: model problems and algorithms, *Nonlinearity* **17** (2004) R55-R127.

[68] Brandt, A., Multiscale solvers and systematic upscaling in computational physics, *Computer Physics Communication*, **169** (2005) 438–441.

[69] Brandt, A., Methods of Systematic Upscaling, Report MCS06-05, Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel 2006

[70] Zlochin, M. and Brandt, A., Systematic Upscaling for Feynman Path Integrals, Progress Report MCS05-10, Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel, 2005