DYNAMICS OF OBSERVABLES IN RANK-BASED MODELS AND PERFORMANCE OF FUNCTIONALLY GENERATED PORTFOLIOS

BY SERGIO A. ALMADA MONTER[*], MYKHAYLO SHKOLNIKOV[†,1] AND JIACHENG ZHANG[†]

JPMorgan Chase[*] and Princeton University[†]

In the seminal work (*Stochastic Portfolio Theory: Stochastic Modelling and Applied Probability* (2002) Springer), several macroscopic market observables have been introduced, in an attempt to find characteristics capturing the diversity of a financial market. Despite the crucial importance of such observables for investment decisions, a concise mathematical description of their dynamics has been missing. We fill this gap in the setting of rank-based models. The results are then used to study the performance of multiplicatively and additively functionally generated portfolios.

1. Introduction. A key characteristic of an equity market is its diversity which, on an intuitive level, describes how evenly the investors distribute their capital among the publicly traded companies. In the seminal work [10] (see also [11–13]), Fernholz has initiated the program of capturing the concept of diversity mathematically, and thus, quantifying its implications on the performance of investment portfolios. Given the market weights $\mu_1(t), \mu_2(t), \ldots, \mu_n(t)$ at a time $t \geq 0$ (i.e., the fractions of market capital invested in the n publicly traded companies at that time), he suggested to measure the market diversity by

$$(1.1) \qquad D_p(t) := \left(\sum_{i=1}^{n} \mu_i(t)^p \right)^{1/p} \quad \text{for some } p \in (0, 1)$$

(see [10], Example 3.4.4). The choice $p \in (0, 1)$ ensures that the right-hand side of (1.1) is a concave function of $\mu_1(t), \mu_2(t), \ldots, \mu_n(t)$ and attains its maximum for the uniform capital distribution $\mu_1(t) = \mu_2(t) = \cdots = \mu_n(t) = \frac{1}{n}$. The limiting case

$$(1.2) \qquad H(t) := \lim_{p \uparrow 1} \frac{p}{1-p} \log D_p(t) = -\sum_{i=1}^{n} \mu_i(t) \log \mu_i(t),$$

known as the market entropy, retains the latter two properties and can therefore be regarded as an alternative measure of the market diversity (cf. [10], Section 2.3). We refer to [10], Figures 6.7, 7.3 and 6.2, for plots of the process $D_{1/2}(\cdot)$ for the largest 1000 companies in the U.S., the process $D_{0.76}(\cdot)$ for the companies forming the S&P 500 index and the process $H(\cdot)$ for the companies in the Center for Research in Securities Prices (CRSP) database of major U.S. stock exchanges (including the NYSE, the AMEX and the NASDAQ), respectively.

Despite the considerable interest in the quantities $D_p(\cdot)$, $p \in (0, 1)$ and $H(\cdot)$ (and the associated functionally generated portfolios, see below), a concise mathematical description of their dynamics has been missing so far. The challenge lies thereby in the fact that, while the vector of market weights $(\mu_1(\cdot), \mu_2(\cdot), \ldots, \mu_n(\cdot))$ is typically modeled by a Markov process, the Markov property is generally not inherited by $D_p(\cdot)$, $p \in (0, 1)$ or $H(\cdot)$. Our first main goal in this paper is to capture the dynamics of nonlinear macroscopic observables of the point process of logarithmic market capitalizations, such as $D_p(\cdot)$ and $H(\cdot)$, in the context of rank-based (a.k.a. first-order) models and when the number of companies n is large. We choose to work with rank-based models because they are known to form the simplest class of market models that is able to reproduce the true long-term average capital distribution of a financial market (see [10], Chapter 5, [2] and [11], Chapter 13). We point to [10], Figure 5.1, for a plot of the latter for the stocks in the CRSP database.

In the context of stochastic portfolio theory, the term *rank-based model* refers to modeling the market weights process $\mu(\cdot) = (\mu_1(\cdot), \mu_2(\cdot), \ldots, \mu_n(\cdot))$ via

$$\left(\frac{e^{X_1^{(n)}(\cdot)}}{\sum_{i=1}^n e^{X_i^{(n)}(\cdot)}}, \frac{e^{X_2^{(n)}(\cdot)}}{\sum_{i=1}^n e^{X_i^{(n)}(\cdot)}}, \ldots, \frac{e^{X_n^{(n)}(\cdot)}}{\sum_{i=1}^n e^{X_i^{(n)}(\cdot)}} \right),$$

where the processes $X_i^{(n)}$, $i = 1, 2, \ldots, n$ form the unique weak solution of the system of stochastic differential equations (SDEs)

(1.3)
$$dX_i^{(n)}(t) = b\big(F_{\varrho^{(n)}(t)}(X_i^{(n)}(t))\big) dt + \sigma\big(F_{\varrho^{(n)}(t)}(X_i^{(n)}(t))\big) dB_i^{(n)}(t),$$
$$i = 1, 2, \ldots, n,$$

with coefficient functions $b : [0, 1] \to \mathbb{R}$ and $\sigma : [0, 1] \to (0, \infty)$, the empirical cumulative distribution functions $F_{\varrho^{(n)}(t)}(x) := \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i^{(n)}(t) \le x\}}$ and independent standard Brownian motions $B_1^{(n)}, B_2^{(n)}, \ldots, B_n^{(n)}$. The system (1.3) is a special case of the systems of SDEs studied in [3]; in particular, the main result of [3] shows the weak uniqueness for (1.3) (the weak existence for (1.3) falls under the classical result of [29], Exercise 12.4.3). More recently, the interacting particle system described by (1.3) and its variants have attracted much attention due to their appearance in stochastic portfolio theory and an open problem of Aldous [1] (see [10], Section 5.5, [2], [11], Section 13, [16–19, 26, 28] for the former and [6, 7, 9, 24, 25, 27, 31] for the latter).

The $n \to \infty$ asymptotics of nonlinear macroscopic observables that we derive herein rely on the law of large numbers for rank-based models in [18], Corollary 2.13 (see also [8], Corollary 1.6, [28], Theorem 1.2) and the associated central limit theorem in [22], Theorem 1.2. Both of these results hold under the following (stronger than the original) assumption.

ASSUMPTION 1.1. (a) There exists a probability measure λ on $\mathbb{R}$ possessing a bounded density function and satisfying

$$(1.4) \qquad \forall \theta > 0 : \quad \int_{\mathbb{R}} e^{\theta |x|} \lambda(\mathrm{d}x) < \infty$$

such that the initial locations of the particles $X_1^{(n)}(0), X_2^{(n)}(0), \ldots, X_n^{(n)}(0)$ are i.i.d. according to λ for all $n \in \mathbb{N}$.

(b) The coefficient functions b and σ are differentiable with locally Hölder continuous derivatives.

We now state the versions of [18], Corollary 2.13 and [22], Theorem 1.2 used in this paper for future reference. Hereby, we write $M_1(\mathbb{R})$ for the space of probability measures on $\mathbb{R}$ equipped with the topology of weak convergence, $C([0, \infty), M_1(\mathbb{R}))$ for the space of continuous functions from $[0, \infty)$ to $M_1(\mathbb{R})$ endowed with the topology of locally uniform convergence, as well as $M_{\mathrm{fin}}(\mathbb{R})$ and $M_{\mathrm{fin}}([0, t] \times \mathbb{R})$, $t > 0$ for the spaces of finite signed measures on $\mathbb{R}$ and $[0, t] \times \mathbb{R}$, $t > 0$ viewed as the duals of $C_0(\mathbb{R})$ and $C_0([0, t] \times \mathbb{R})$, $t > 0$ with the associated weak-$*$ topologies, respectively.

PROPOSITION 1.2 (cf. [18], Corollary 2.13). *Under Assumption 1.1, the processes of empirical measures $\varrho^{(n)}(\cdot) = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i^{(n)}(\cdot)}$, $n \in \mathbb{N}$ converge in probability in $C([0, \infty), M_1(\mathbb{R}))$ to a deterministic limit $\varrho(\cdot)$, where the corresponding process of cumulative distribution functions $R(t, \cdot) := F_{\varrho(t)}(\cdot)$, $t \geq 0$ is the unique generalized solution of the Cauchy problem for the porous medium equation*

$$(1.5) \qquad R_t = -B(R)_x + \Sigma(R)_{xx}, \qquad R(0, \cdot) = F_\lambda(\cdot)$$

in the sense of [14], Definition 3, $B(r) := \int_0^r b(a)\,\mathrm{d}a$, $\Sigma(r) := \int_0^r \frac{1}{2}\sigma(a)^2\,\mathrm{d}a$, and $F_\lambda(\cdot)$ is the cumulative distribution function of λ.

PROPOSITION 1.3 (cf. [22], Theorem 1.2). *Let Assumption 1.1 be satisfied and G be the mild solution of the stochastic partial differential equation (SPDE)*

$$(1.6) \qquad \begin{aligned} G_t &= -\big(b(R)G\big)_x + \left(\frac{\sigma(R)^2}{2}G\right)_{xx} + \sigma(R)R_x^{1/2}\dot{W}, \\ G(0, \cdot) &= \beta\big(F_\lambda(\cdot)\big), \end{aligned}$$

with the function R from Proposition 1.2, the space-time white noise $\dot{W}$, and a standard Brownian bridge β independent of $\dot{W}$. Then the sequences of processes

(1.7)
$$t \mapsto \sqrt{n}\big(F_{\varrho^{(n)}(t)}(x) - R(t, x)\big)\,\mathrm{d}x, \quad n \in \mathbb{N} \quad and$$
$$t \mapsto \sqrt{n}\big(F_{\varrho^{(n)}(s)}(x) - R(s, x)\big)\,\mathrm{d}x\,\mathrm{d}s, \quad n \in \mathbb{N},$$

with values in $M_{\mathrm{fin}}(\mathbb{R})$ and $M_{\mathrm{fin}}([0, t] \times \mathbb{R}), t > 0$, respectively, converge jointly in the finite-dimensional distribution sense to

(1.8)
$$t \mapsto G(t, x)\,\mathrm{d}x \quad and \quad t \mapsto G(s, x)\mathbf{1}_{[0,t] \times \mathbb{R}}(s, x)\,\mathrm{d}s\,\mathrm{d}x.$$

We are now ready to give the first two main results of the present work, which yield a comprehensive description of the large n asymptotic dynamics for nonlinear macroscopic observables of the form

(1.9)
$$\mathcal{J}_{J; f_1, \dots, f_k}(\alpha(\cdot)) := J\left(\int_{\mathbb{R}} f_1\,\mathrm{d}\alpha(\cdot), \dots, \int_{\mathbb{R}} f_k\,\mathrm{d}\alpha(\cdot)\right),$$

where $\alpha(\cdot) \in C([0, \infty), M_1(\mathbb{R}))$, J is a continuously differentiable function, and

(1.10)
$$f_1, \dots, f_k \in \mathcal{E}_\ell$$
$$:= \left\{ f \in C^\ell(\mathbb{R}) : \left|\frac{\mathrm{d}^\ell f}{\mathrm{d}x^\ell}(x)\right| \le C e^{C|x|}, x \in \mathbb{R} \text{ for some } C \ge 0 \right\},$$

with $\ell = 1$ in Theorem 1.4 and $\ell = 3$ in Theorem 1.5 below. For simplicity, we use henceforth the bilinear form notation $\langle f, \nu \rangle$ for $\int_{\mathbb{R}} f\,\mathrm{d}\nu$ and write $(f_1, \dots, f_k) \in \mathcal{E}_\ell^U$ for a $U \subseteq \mathbb{R}^k$ whenever $f_1, \dots, f_k \in \mathcal{E}_\ell$ and

$$\big(\langle f_1, \nu \rangle, \dots, \langle f_k, \nu \rangle\big) \in U$$

$$\text{for all } \nu \in M_1(\mathbb{R}) \text{ fulfilling } \int_{\mathbb{R}} e^{\theta|x|}\nu(\mathrm{d}x) < \infty, \theta > 0.$$

THEOREM 1.4. *Let Assumption 1.1 be satisfied and $(f_1, \dots, f_k) \in \mathcal{E}_1^U$ for a convex open $U \subseteq \mathbb{R}^k$. Then, for all $J \in C^1(U)$, one has*

(1.11)
$$\sqrt{n}\big(\mathcal{J}_{J; f_1, \dots, f_k}(\varrho^{(n)}(\cdot)) - \mathcal{J}_{J; f_1, \dots, f_k}(\varrho(\cdot))\big)$$
$$\overset{n \to \infty}{\longrightarrow} -\sum_{j=1}^{k} \mathcal{J}_{J_{x_j}; f_1, \dots, f_k}(\varrho(\cdot)) \int_{\mathbb{R}} f_j'(x)G(\cdot, x)\,\mathrm{d}x$$

in the finite-dimensional distribution sense.

Theorem 1.4 provides the fundamental tool in establishing the central limit theorem for the hitting times of nonlinear macroscopic observables.

THEOREM 1.5. *Let Assumption* 1.1 *be satisfied,* $(f_1, \ldots, f_k) \in \mathcal{E}_3^U$ *for a convex open* $U \subseteq \mathbb{R}^k$, *and* $J \in C^1(U)$. *Suppose* $a \in \mathbb{R}$ *is such that*

$$\tau := \inf\{t \geq 0 : \mathcal{J}_{J; f_1, \ldots, f_k}(\varrho(t)) = a\} < \infty \quad and$$

(1.12)
$$\frac{d\mathcal{J}_{J; f_1, \ldots, f_k}(\varrho(\cdot))}{dt}(\tau) \neq 0.$$

Then the sequence of hitting times

(1.13) $$\tau^{(n)} := \inf\{t \geq 0 : \mathcal{J}_{J; f_1, \ldots, f_k}(\varrho^{(n)}(t)) = a\}, \quad n \in \mathbb{N}$$

converges in distribution when properly rescaled:

(1.14) $$\sqrt{n}(\tau^{(n)} - \tau) \xrightarrow{n \to \infty} \frac{\sum_{j=1}^{k} \mathcal{J}_{J_{x_j}; f_1, \ldots, f_k}(\varrho(\tau)) \int_{\mathbb{R}} f_j'(x) G(\tau, x) \, dx}{\frac{d\mathcal{J}_{J; f_1, \ldots, f_k}(\varrho(\cdot))}{dt}(\tau)}.$$

REMARK 1.6. We emphasize that Theorems 1.4 and 1.5 apply, in particular, to the (appropriately normalized) processes $D_p(\cdot)$, $p \in (0, 1)$ and $H(\cdot)$ of (1.1) and (1.2), respectively, allowing to approximate them by Gaussian processes (Theorem 1.4) and their hitting times by Gaussian random variables (Theorem 1.5). For the exact statements, please see Corollary 5.3 and Remark 5.5 below.

REMARK 1.7. The condition of finiteness of all exponential moments on λ in Assumption 1.1(a) enters naturally in the context of market observables from stochastic portfolio theory, which often (e.g., in the case of $D_p(\cdot)$, $p \in (0, 1)$) involve powers of the market capitalizations that, in turn, are images of $X_1^{(n)}(\cdot), X_2^{(n)}(\cdot), \ldots, X_n^{(n)}(\cdot)$ under the exponential function.

1.1. *Application to functionally generated portfolios.* Theorems 1.4 and 1.5 can be used to obtain concentration of measure estimates on the performance of functionally generated portfolios, both additive and multiplicative, as defined in [10], Chapter 3 and [20], respectively. For functionally generated portfolios, the portfolio weights process $\pi(\cdot) = (\pi_1(\cdot), \pi_2(\cdot), \ldots, \pi_n(\cdot))$, which records the fractions of capital invested in the n companies over time, is given in terms of a function $\widetilde{\Psi}$ of the market weights and its derivatives. The crucial observation in [10], Chapter 3 and [20] is that, for two carefully constructed classes of portfolio weights processes $\pi(\cdot)$, the value $\frac{X^\pi(\cdot)}{X^\mu(\cdot)}$ of $\pi(\cdot)$ relative to that of the market portfolio $\mu(\cdot)$, with

(1.15)
$$\frac{dX^\pi(\cdot)}{X^\pi(\cdot)} = \sum_{i=1}^{n} \pi_i(\cdot) \frac{dX_i^{(n)}(\cdot)}{X_i^{(n)}(\cdot)},$$

$$\frac{dX^\mu(\cdot)}{X^\mu(\cdot)} = \sum_{i=1}^{n} \mu_i(\cdot) \frac{dX_i^{(n)}(\cdot)}{X_i^{(n)}(\cdot)} \quad \text{and} \quad X^\pi(0) = X^\mu(0),$$

admits a pathwise representation. Let us be more specific in this direction.

For a twice continuously differentiable function $\widetilde{\Psi}$ on an open neighborhood of the unit simplex $\{x \in (0, 1)^n : \sum_{i=1}^n x_i = 1\} \subset \mathbb{R}^n$, one formally defines the portfolio weights processes $\pi^{\widetilde{\Psi};\times}(\cdot)$ and $\pi^{\widetilde{\Psi};+}(\cdot)$ multiplicatively and additively generated by $\widetilde{\Psi}$ according to

$$\pi_i^{\widetilde{\Psi};\times}(\cdot) = \left((\log \widetilde{\Psi})_{x_i}(\mu(\cdot)) + 1 - \sum_{j=1}^n \mu_j(\cdot)(\log \widetilde{\Psi})_{x_j}(\mu(\cdot)) \right) \mu_i(\cdot),$$

$$i = 1, 2, \ldots, n,$$

$$\pi_i^{\widetilde{\Psi};+}(\cdot) = \left(\frac{\widetilde{\Psi}_{x_i}(\mu(\cdot)) - \sum_{j=1}^n \mu_j(\cdot)\widetilde{\Psi}_{x_j}(\mu(\cdot))}{\widetilde{\Psi}(\mu(\cdot)) - \frac{1}{2} \sum_{i,j=1}^n \int_0^{\cdot} \widetilde{\Psi}_{x_i x_j}(\mu(s)) \, d[\mu_i, \mu_j](s)} + 1 \right) \mu_i(\cdot),$$

$$i = 1, 2, \ldots, n,$$

respectively. The key feature of the functionally generated portfolios $\pi^{\widetilde{\Psi};\times}(\cdot)$, $\pi^{\widetilde{\Psi};+}(\cdot)$ is that their values $V^{\widetilde{\Psi};\times}(\cdot)$, $V^{\widetilde{\Psi};+}(\cdot)$ relative to that of the market portfolio $\mu(\cdot)$ admit pathwise representations, which under the usual convention $V^{\widetilde{\Psi};\times}(0) = V^{\widetilde{\Psi};+}(0) = 1$ read

$$(1.16) \quad V^{\widetilde{\Psi};\times}(t) = \frac{\widetilde{\Psi}(\mu(t))}{\widetilde{\Psi}(\mu(0))} \exp\left(-\frac{1}{2} \sum_{i,j=1}^n \int_0^t \frac{\widetilde{\Psi}_{x_i x_j}(\mu(\cdot))}{\widetilde{\Psi}(\mu(\cdot))} \, d[\mu_i, \mu_j](\cdot) \right),$$

$$t \geq 0,$$

$$V^{\widetilde{\Psi};+}(t) = 1 + \widetilde{\Psi}(\mu(t)) - \widetilde{\Psi}(\mu(0))$$

$$(1.17) \quad\quad\quad - \frac{1}{2} \sum_{i,j=1}^n \int_0^t \widetilde{\Psi}_{x_i x_j}(\mu(\cdot)) \, d[\mu_i, \mu_j](\cdot), \quad t \geq 0$$

(cf. [11], equation (11.2), [20], equation (4.3)).

The function $\widetilde{\Psi}$ usually corresponds to a function Ψ of the form

$$(1.18) \quad \Psi : \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto J\left(\frac{1}{n} \sum_{i=1}^n f_1(x_i), \ldots, \frac{1}{n} \sum_{i=1}^n f_k(x_i) \right)$$

and with the homogeneity property

$$(1.19) \quad \forall x \in \mathbb{R}^n, r \in \mathbb{R}: \quad \Psi(x) = \Psi(x_1 + r, x_2 + r, \ldots, x_n + r)$$

via the relation

$$(1.20) \quad \Psi(x) = \widetilde{\Psi}\left(\frac{e^{x_1}}{\sum_{i=1}^n e^{x_i}}, \frac{e^{x_2}}{\sum_{i=1}^n e^{x_i}}, \ldots, \frac{e^{x_n}}{\sum_{i=1}^n e^{x_i}} \right), \quad x \in \mathbb{R}^n.$$

We assume henceforth that $\widetilde{\Psi}$ is positive and concave in the setting of (1.16) or concave in the setting of (1.17), since then the respective *excess growth processes*,

defined as

$$\Gamma^{\widetilde{\Psi};\times}(t) = -\frac{1}{2} \sum_{i,j=1}^{n} \int_0^t \frac{\widetilde{\Psi}_{x_i x_j}(\mu(\cdot))}{\widetilde{\Psi}(\mu(\cdot))} \, \mathrm{d}[\mu_i, \mu_j](\cdot), \quad t \geq 0,$$

$$\Gamma^{\widetilde{\Psi};+}(t) = -\frac{1}{2} \sum_{i,j=1}^{n} \int_0^t \widetilde{\Psi}_{x_i x_j}(\mu(\cdot)) \, \mathrm{d}[\mu_i, \mu_j](\cdot), \quad t \geq 0,$$

are nondecreasing (see [20], Example 3.5). In particular, by (1.16), (1.17) the associated value processes $V^{\widetilde{\Psi};\times}(\cdot)$ or $V^{\widetilde{\Psi};+}(\cdot)$ reach levels $v > 1$ before the hitting times $\tau^{(n)}$ of levels $a = v\widetilde{\Psi}(\mu(0))$ or $a = v - 1 + \widetilde{\Psi}(\mu(0))$, respectively, whose asymptotics are described by Theorem 1.5.

More precise estimates on the processes $V^{\widetilde{\Psi};\times}(\cdot)$ and $V^{\widetilde{\Psi};+}(\cdot)$ can be obtained under the additional assumptions

$$\forall n \in \mathbb{N}, i = 1, 2, \ldots, n - 1:$$

(1.21)
$$\frac{1}{i} \sum_{j=1}^{i} b\left(\frac{j}{n}\right) > \frac{1}{n-i} \sum_{j=i+1}^{n} b\left(\frac{j}{n}\right) \quad \text{and} \quad \sigma(\cdot) = 1$$

(see [25], Remark on page 2187, for a detailed discussion of the first assumption; in addition, note that the constant 1 in the second assumption can be turned into any other positive constant by a deterministic time change). Indeed, under the assumptions in (1.21), [16], Corollary 8, applies and can be naturally combined with Theorem 1.5. In order to avoid cumbersome notation, we state the next corollary for the multiplicative generation only and refer to Remark 1.10 for a discussion of the additive case.

COROLLARY 1.8. *Let Assumption* 1.1, *the assumptions in* (1.21), *and for all* $n \in \mathbb{N}$,

(1.22)
$$\operatorname*{ess\,sup}_{\omega,t} \sum_{i,j=1}^{n} \frac{\widetilde{\Psi}_{x_i x_j}(\mu(\cdot))}{\widetilde{\Psi}(\mu(\cdot))} \frac{\mathrm{d}[\mu_i, \mu_j](\cdot)}{\mathrm{d}t}$$

$$- \operatorname*{ess\,inf}_{\omega,t} \sum_{i,j=1}^{n} \frac{\widetilde{\Psi}_{x_i x_j}(\mu(\cdot))}{\widetilde{\Psi}(\mu(\cdot))} \frac{\mathrm{d}[\mu_i, \mu_j](\cdot)}{\mathrm{d}t} \in (0, \infty)$$

be satisfied. Then, for the functions J *and* $f_1, f_2, \ldots, f_k$ *of* (1.18), (1.20), a, τ *as in* (1.12), $r^{\times} := -\lim_{t \to \infty} \frac{1}{2t} \int_0^t \frac{\widetilde{\Psi}_{x_i x_j}(\mu(\cdot))}{\widetilde{\Psi}(\mu(\cdot))} \, \mathrm{d}[\mu_i, \mu_j](\cdot)$ *and* $r, s > 0$, *the stopping time*

(1.23)
$$\eta^{\widetilde{\Psi};\times} := \inf\left\{ t \geq 0 : V^{\widetilde{\Psi};\times}(t) = \frac{a}{\widetilde{\Psi}(\mu(0))} e^{(r^{\times} - r)(\tau - s/\sqrt{n})} \right\}$$

obeys for all $\varepsilon > 0$ the estimate

$$
\mathbb{P}\big(\eta^{\widetilde{\Psi};\times} \geq \tau + s/\sqrt{n}\big) \leq 2\overline{\Phi}(s/\chi)\big(1 + o_n(1)\big)
$$
(1.24)
$$
+ \left\| \frac{\mathrm{d}\kappa^{(n)}}{\mathrm{d}\zeta^{(n)}} \right\|_{L^2(\zeta^{(n)})} e^{-c^{\times}(r,\varepsilon)(\tau - s/\sqrt{n})}.
$$

Hereby, $\overline{\Phi}$ is the standard normal tail cumulative distribution function; χ is the standard deviation of the random variable on the right-hand side of (1.14); *$o_n(1)$ is a quantity tending to 0 as $n \to \infty$; $\kappa^{(n)}$ and $\zeta^{(n)}$ are the laws of the vector of differences between the consecutive order statistics of $(X_1^{(n)}, X_2^{(n)}, \ldots, X_n^{(n)})$ at time 0 and in stationarity, respectively; and*

$$
c^{\times}(r, \varepsilon)
$$

$$
= \frac{\min_{1 \leq j \leq n-1}\big(\sum_{i=1}^{j} b(\frac{i}{n}) - \frac{j}{n}\sum_{i=1}^{n} b(\frac{i}{n})\big)^2}{2 - 2\cos\frac{\pi}{n}}
$$
(1.25)
$$
\cdot \max\bigg(\frac{r^2}{(C^{\times})^2},
$$
$$
4\varepsilon\big(\varepsilon + v^{\times}\big)\bigg(\sqrt{1 + \frac{r^2}{2\varepsilon(\varepsilon + v^{\times})^2 \max(|C^{\times,\uparrow}|, |C^{\times,\downarrow}|)^2}} - 1\bigg)\bigg),
$$

with $C^{\times}$, $C^{\times,\uparrow}$, $C^{\times,\downarrow}$, and $v^{\times}$ being the overall expression, the essential supremum, the essential infimum and the variance under $\zeta^{(n)}$ of the expression inside the essential supremum in (1.22).

REMARK 1.9. Inequality (1.24) can be interpreted as follows. If one invests in the portfolio generated multiplicatively by a function $\widetilde{\Psi}$ satisfying the condition (1.22) and aims for the associated process $\widetilde{\Psi}(\mu(\cdot))$ to reach an admissible value of a (i.e., one for which (1.12) holds), then one will achieve a logarithmic return relative to the market portfolio $\mu(\cdot)$ of $\log a - \log \widetilde{\Psi}(\mu(0)) + (r^{\times} - r)(\tau - s/\sqrt{n})$ before time $\tau + s/\sqrt{n}$ with a confidence probability of at least one minus the right-hand side of (1.24). We note that, in the practically relevant regime $a \geq \mathcal{J}_{J;f_1,\ldots,f_k}(\varrho(0))$ and $r^{\times} \geq r$, by increasing the values of a and s one can increase both the relative return and the confidence probability for sufficiently large n, at the expense of thereby increasing the upper bound $\tau + s/\sqrt{n}$ on the investment horizon.

REMARK 1.10. In the case of additive generation, a concentration estimate similar to (1.24) holds for

$$
\eta^{\widetilde{\Psi};+} := \inf\{t \geq 0 : V^{\widetilde{\Psi};+}(t) = 1 + a - \widetilde{\Psi}(\mu(0)) + (r^+ - r)(\tau - s/\sqrt{n})\},
$$

assuming

$$
\operatorname*{ess\,sup}_{\omega,t} \sum_{i,j=1}^{n} \widetilde{\Psi}_{x_i x_j}(\mu(\cdot)) \frac{\mathrm{d}[\mu_i, \mu_j](\cdot)}{\mathrm{d}t}
$$

(1.26)

$$
- \operatorname*{ess\,inf}_{\omega,t} \sum_{i,j=1}^{n} \widetilde{\Psi}_{x_i x_j}(\mu(\cdot)) \frac{\mathrm{d}[\mu_i, \mu_j](\cdot)}{\mathrm{d}t} \in (0, \infty)
$$

instead of (1.22) and with $r^+ := -\lim_{t\to\infty} \frac{1}{2t} \sum_{i,j=1}^{n} \int_0^t \widetilde{\Psi}_{x_i x_j}(\mu(\cdot)) \, \mathrm{d}[\mu_i, \mu_j](\cdot)$ taking the place of $r^\times$. We leave the straightforward adaptation to this case of the proof of Corollary 1.8, given Proposition 5.9 below, to the reader.

The rest of the paper is structured as follows. In Section 2, we collect some results from [15, 22] and [4] that are used repeatedly in the proofs of Theorems 1.4 and 1.5. Section 3 is then devoted to the proof of Theorem 1.4. The latter is based on Proposition 1.3, but requires significant additional work due to the exponential growth at infinity of the derivatives of functions in $\mathcal{E}_1$ and the nonlinearity of J. In particular, the proof invokes the mean stochastic comparison of [15] and the quantitative propagation of chaos result of [22], Theorem 1.6. Subsequently, we give the proof of Theorem 1.5 in Section 4, which relies on the previously mentioned tools and a creative reduction to the estimate on the expected Wasserstein distance W_1 between the empirical measure of an i.i.d. sample and the underlying distribution in [4], Theorem 3.2. Next, in Section 5, we apply Theorems 1.4 and 1.5 to the main examples of diversity measures from stochastic portfolio theory. Lastly, in Section 5.3, we provide the proof of Corollary 1.8.

2. Preliminaries. The general *propagation of chaos* paradigm (see [30]) suggests that under Assumption 1.1, for large values of n, the weak solution of (1.3) should be well approximated by the strong solution of

(2.1)
$$
d\overline{X}_i^{(n)}(t) = b\big(R\big(t, \overline{X}_i^{(n)}(t)\big)\big) \, \mathrm{d}t + \sigma\big(R\big(t, \overline{X}_i^{(n)}(t)\big)\big) \, \mathrm{d}B_i^{(n)}(t),
$$
$$
i = 1, 2, \ldots, n
$$

with the initial condition $\overline{X}_i^{(n)}(0) = X_i^{(n)}(0)$, $i = 1, 2, \ldots, n$, the function R from Proposition 1.2 and the same Brownian motions $B_1^{(n)}, B_2^{(n)}, \ldots, B_n^{(n)}$ as in (1.3). Indeed, under Assumption 1.1, the coefficient functions $(t, x) \mapsto b(R(t, x))$, $(t, x) \mapsto \sigma(R(t, x))$ are uniformly Lipschitz in x on any strip of the form $[0, T] \times \mathbb{R}$ by [22], Proposition 2.5, so that the strong existence and uniqueness for (2.1) readily follow. The propagation of chaos estimates from [22], Theorem 1.6, quantify the term "well approximated."

PROPOSITION 2.1 (cf. [22], Theorem 1.6).   *Let Assumption* 1.1 *be satisfied. Then, for all* $p, T > 0$, *one can find a constant* $C = C(p, T) < \infty$ *such that*

$$(2.2) \qquad \mathbb{E}\Big[\sup_{t \in [0,T]} |X_i^{(n)}(t) - \overline{X}_i^{(n)}(t)|^p \Big] \leq C n^{-p/2}, \quad i = 1, 2, \ldots, n, n \in \mathbb{N}.$$

Next, we introduce the process of empirical measures

$$(2.3) \qquad \overline{\varrho}^{(n)}(\cdot) = \frac{1}{n} \sum_{i=1}^{n} \delta_{\overline{X}_i^{(n)}(\cdot)}$$

and recall that the Wasserstein distance of order one is defined for any $\nu_1, \nu_2 \in M_1(\mathbb{R})$ with finite first moments by

$$(2.4) \qquad W_1(\nu_1, \nu_2) = \inf_{Z_1 \overset{d}{=} \nu_1, Z_2 \overset{d}{=} \nu_2} \mathbb{E}[|Z_1 - Z_2|].$$

Each $\overline{\varrho}^{(n)}(t)$ constitutes the empirical measure of an i.i.d. sample from the probability measure $\varrho(t)$ introduced in Proposition 1.2. Hence, we may aim to bound the associated expected W_1-distance $\mathbb{E}[W_1(\overline{\varrho}^{(n)}(t), \varrho(t))]$ by means of [4], Theorem 3.2, which requires a moment estimate for $\varrho(t)$. The latter, in turn, can be obtained under Assumption 1.1 from the mean stochastic comparison results of [15] as follows. With $C_b^{\uparrow} := \max_{a \in [0,1]} b(a)$, $C_b^{\downarrow} := \min_{a \in [0,1]} b(a)$ and $C_\sigma := \max_{a \in [0,1]} |\sigma(a)|$, consider the Brownian motions with drift:

$$(2.5) \qquad dY^{\uparrow}(t) = C_b^{\uparrow} \, dt + C_\sigma \, dB_1^{(1)}(t), \quad Y^{\uparrow}(0) \overset{d}{=} \lambda,$$

$$(2.6) \qquad dY^{\downarrow}(t) = C_b^{\downarrow} \, dt + C_\sigma \, dB_1^{(1)}(t), \quad Y^{\downarrow}(0) \overset{d}{=} \lambda.$$

The next proposition is then a direct consequence of [15], inequality (1.5) and page 318, Remark (4).

PROPOSITION 2.2.   *Let Assumption* 1.1 *be satisfied. Then, for all* $i = 1, 2, \ldots, n, n \in \mathbb{N}, T > 0, M \in \mathbb{R}$ *and* $\theta > 0$, *one has the comparison results*

$$(2.7) \qquad \begin{aligned} &\mathbb{P}\Big(\sup_{t \in [0,T]} |X_i^{(n)}(t)| \geq M \Big) \vee \mathbb{P}\Big(\sup_{t \in [0,T]} |\overline{X}_i^{(n)}(t)| \geq M \Big) \\ &\leq 2\mathbb{P}\Big(\sup_{t \in [0,T]} Y^{\uparrow}(t) \geq M \Big) + 2\mathbb{P}\Big(\sup_{t \in [0,T]} (-Y^{\downarrow}(t)) \geq M \Big), \end{aligned}$$

$$(2.8) \qquad \begin{aligned} &\mathbb{E}\Big[\sup_{t \in [0,T]} e^{\theta |X_i^{(n)}(t)|} \Big] \vee \mathbb{E}\Big[\sup_{t \in [0,T]} e^{\theta |\overline{X}_i^{(n)}(t)|} \Big] \\ &\leq \mathbb{E}\Big[\sup_{t \in [0,T]} e^{\theta Y^{\uparrow}(t)} \Big] + \mathbb{E}\Big[\sup_{t \in [0,T]} e^{-\theta Y^{\downarrow}(t)} \Big] < \infty. \end{aligned}$$

In particular, the bound of (2.8) allows us to use [4], Theorem 3.2, to estimate each of the quantities $\mathbb{E}[W_1(\overline{\varrho}^{(n)}(t), \varrho(t))]$. Hereby, we keep in mind the alternative representation of the W_1-distance as the L^1-distance between the cumulative distribution functions (see, e.g., [4], Theorem 2.9):

$$(2.9) \qquad W_1(\nu_1, \nu_2) = \int_{\mathbb{R}} |F_{\nu_1}(x) - F_{\nu_2}(x)| \, \mathrm{d}x.$$

PROPOSITION 2.3. *Let Assumption* 1.1 *be satisfied. Then, for all* $T > 0$, *one can find a constant* $C = C(T) < \infty$ *such that*

$$(2.10) \qquad \sup_{t \in [0,T]} \mathbb{E}[W_1(\overline{\varrho}^{(n)}(t), \varrho(t))] \le C n^{-1/2}, \quad n \in \mathbb{N}.$$

3. Proof of Theorem 1.4. Our starting point for the proof of Theorem 1.4 is the identity

$$
\begin{aligned}
(3.1) \qquad &\sqrt{n}\big(\mathcal{J}_{J; f_1, \dots, f_k}(\varrho^{(n)}(\cdot)) - \mathcal{J}_{J; f_1, \dots, f_k}(\varrho(\cdot))\big) \\
&= \sqrt{n}\big(\langle f_1, \varrho^{(n)}(\cdot) - \varrho(\cdot)\rangle, \dots, \langle f_k, \varrho^{(n)}(\cdot) - \varrho(\cdot)\rangle\big) \\
&\quad \times \nabla J\big(\langle f_1, \widetilde{\varrho}^{(n)}(\cdot)\rangle, \dots, \langle f_k, \widetilde{\varrho}^{(n)}(\cdot)\rangle\big)
\end{aligned}
$$

due to the mean value theorem, where $\widetilde{\varrho}^{(n)}(\cdot) = \xi^{(n)}(\cdot)\varrho^{(n)}(\cdot) + (1 - \xi^{(n)}(\cdot))\varrho(\cdot)$ and $\xi^{(n)}(\cdot)$ can be chosen as stochastic processes with values in $M_1(\mathbb{R})$ and $[0, 1]$, respectively, by the Borel selection result of [5], Theorem 6.9.6. The proof of Theorem 1.4 is carried out by studying the convergence of the vector-valued stochastic processes

$$(3.2) \qquad I_1^{(n)}(\cdot) := \sqrt{n}\big(\langle f_1, \varrho^{(n)}(\cdot) - \varrho(\cdot)\rangle, \dots, \langle f_k, \varrho^{(n)}(\cdot) - \varrho(\cdot)\rangle\big),$$

$$(3.3) \qquad I_2^{(n)}(\cdot) := \big(\langle f_1, \widetilde{\varrho}^{(n)}(\cdot)\rangle, \dots, \langle f_k, \widetilde{\varrho}^{(n)}(\cdot)\rangle\big)$$

as $n \to \infty$. In both cases, it is helpful to introduce, for each $M > 0$, an auxiliary function $h_M \in C^\infty(\mathbb{R})$ with values in $[0, 1]$ such that $h_M(x) = 1$ if $|x| \le M$, $h_M(x) = 0$ if $|x| > M + 1$, and

$$(3.4) \qquad \sup_{M > 0} \sup_{x \in \mathbb{R}} |h_M'(x)| \vee \sup_{M > 0} \sup_{x \in \mathbb{R}} |h_M''(x)| < \infty.$$

In addition, we denote $(1 - h_M)$ by $\widehat{h}_M$ for each $M > 0$.

Convergence of $I_1^{(n)}(\cdot)$. With the mild solution G of the SPDE (1.6), we claim that

$$(3.5) \qquad I_1^{(n)}(\cdot) \overset{n \to \infty}{\longrightarrow} -\left(\int_{\mathbb{R}} f_1'(x)G(\cdot, x)\,\mathrm{d}x, \dots, \int_{\mathbb{R}} f_k'(x)G(\cdot, x)\,\mathrm{d}x\right)$$

in the finite-dimensional distribution sense. To this end, we write each component $-\sqrt{n}\langle f_j, \varrho^{(n)}(\cdot) - \varrho(\cdot)\rangle$ of $-I_1^{(n)}(\cdot)$ as

$$
\sqrt{n} \int_{\mathbb{R}} f_j'(x) h_M(x) \big(F_{\rho^{(n)}(\cdot)}(x) - R(\cdot, x)\big)\, dx
$$

$$
+ \sqrt{n} \int_{\mathbb{R}} f_j(x) h_M'(x) \big(F_{\rho^{(n)}(\cdot)}(x) - R(\cdot, x)\big)\, dx
$$

(3.6)

$$
+ \sqrt{n} \int_{\mathbb{R}} f_j'(x) \widehat{h}_M(x) \big(F_{\rho^{(n)}(\cdot)}(x) - R(\cdot, x)\big)\, dx
$$

$$
+ \sqrt{n} \int_{\mathbb{R}} f_j(x) \widehat{h}_M'(x) \big(F_{\rho^{(n)}(\cdot)}(x) - R(\cdot, x)\big)\, dx
$$

using $f_j = f_j h_M + f_j \widehat{h}_M$ and integration by parts (observe that the boundary terms thereby vanish thanks to $f_j \widehat{h}_M \in \mathcal{E}_0$, the estimate (2.8) and Markov's inequality).

Taking first the $n \to \infty$ limit and then the $M \to \infty$ limit of the first summand in (3.6) for $j = 1, 2, \ldots, k$ gives

(3.7)
$$
\left(\int_{\mathbb{R}} f_1'(x) G(\cdot, x)\, dx, \ldots, \int_{\mathbb{R}} f_k'(x) G(\cdot, x)\, dx \right)
$$

in the finite-dimensional distribution sense. Indeed, by Proposition 1.3 the $n \to \infty$ limit results in the mean zero Gaussian process

(3.8)
$$
\left(\int_{\mathbb{R}} f_1'(x) h_M(x) G(\cdot, x)\, dx, \ldots, \int_{\mathbb{R}} f_k'(x) h_M(x) G(\cdot, x)\, dx \right).
$$

For its convergence in finite-dimensional distribution as $M \to \infty$ to the mean zero Gaussian process in (3.7), it suffices to verify the convergence of the corresponding covariance functions. Upon a decomposition of $f_1', f_2', \ldots, f_k'$ into the positive and negative parts, the positivity of the covariance function of G (see [22], Remark 1.4) and the monotone convergence theorem allow to reduce the convergence of the covariance functions to a statement about the uniform boundedness of the variances involved.

LEMMA 3.1. *Let Assumption* 1.1 *be satisfied. Then, for all* $t \geq 0$ *and* $j \in \{1, 2, \ldots, k\}$,

(3.9)
$$
\sup_{M > 0} \mathbb{E}\left[\left(\int_{\mathbb{R}} f_j'(x)_+ h_M(x) G(t, x)\, dx \right)^2 \right]
$$

$$
\vee\, \mathbb{E}\left[\left(\int_{\mathbb{R}} f_j'(x)_- h_M(x) G(t, x)\, dx \right)^2 \right] < \infty.
$$

Assuming Lemma 3.1, the proof of (3.5) hinges on the next lemma, which shows that the contributions of the second, third and fourth summands in (3.6) to the $n \to \infty$ limit of $I_1^{(n)}(\cdot)$ become negligible as M tends to infinity.

LEMMA 3.2. *Let Assumption* 1.1 *be satisfied. Then, for any* $\varepsilon > 0$, $t \geq 0$, $f_0 \in \mathcal{E}_0$ *and uniformly bounded family of functions* $g_M : \mathbb{R} \to \mathbb{R}$, $M > 0$ *such that* $g_M(x) = 0$, $x \in [-M, M]$ *for each* $M > 0$,

$$
\begin{aligned}
(3.10) \quad & \limsup_{M \to \infty} \limsup_{n \to \infty} \mathbb{P}\left(\left| \sqrt{n} \int_{\mathbb{R}} f_0(x) g_M(x) (F_{\varrho^{(n)}(t)}(x) - R(t, x)) \, dx \right| > \varepsilon \right) \\
& = 0.
\end{aligned}
$$

We proceed to the proofs of the two lemmas.

PROOF OF LEMMA 3.1. For all $M > 0$, we have

$$
\begin{aligned}
(3.11) \quad & \mathbb{E}\left[\left(\int_{\mathbb{R}} f_j'(x)_+ h_M(x) G(t, x) \, dx \right)^2 \right] \\
& \leq \liminf_{n \to \infty} \mathbb{E}\left[\left(\int_{\mathbb{R}} f_j'(x)_+ h_M(x) \sqrt{n} (F_{\varrho^{(n)}(t)}(x) - R(t, x)) \, dx \right)^2 \right]
\end{aligned}
$$

by Proposition 1.3, Skorokhod's representation theorem and Fatou's lemma. With $f_{j,M;+}(x) := \int_0^x f_j'(y)_+ h_M(y) \, dy$, integration by parts yields for the term inside the latter limit inferior

$$
\begin{aligned}
(3.12) \quad & n \mathbb{E}[\langle f_{j,M;+}, \varrho^{(n)}(t) - \varrho(t) \rangle^2] \\
& \leq 2n \mathbb{E}[\langle f_{j,M;+}, \varrho^{(n)}(t) - \overline{\varrho}^{(n)}(t) \rangle^2] + 2n \mathbb{E}[\langle f_{j,M;+}, \overline{\varrho}^{(n)}(t) - \varrho(t) \rangle^2].
\end{aligned}
$$

Next, we insert the definitions of $\varrho^{(n)}(t)$, $\overline{\varrho}^{(n)}(t)$, apply the Cauchy–Schwarz inequality and exploit the independence of $\overline{X}_1^{(n)}(t) \stackrel{d}{=} \overline{X}_2^{(n)}(t) \stackrel{d}{=} \cdots \stackrel{d}{=} \overline{X}_n^{(n)}(t) \stackrel{d}{=} \varrho(t)$ to get

$$
\begin{aligned}
(3.13) \quad & 2\mathbb{E}\left[\sum_{i=1}^n (f_{j,M;+}(X_i^{(n)}(t)) - f_{j,M;+}(\overline{X}_i^{(n)}(t)))^2 \right] \\
& \quad + 2\mathbb{E}[(f_{j,M;+}(\overline{X}_1^{(n)}(t)) - \langle f_{j,M;+}, \varrho(t) \rangle)^2].
\end{aligned}
$$

Since $f_1, \ldots, f_k \in \mathcal{E}_1$, we can pick a constant $C < \infty$ independent of j and M such that $|f_{j,M;+}'(x)| \leq C e^{C|x|}$, $x \in \mathbb{R}$ and $|f_{j,M;+}(x)| \leq C e^{C|x|}$, $x \in \mathbb{R}$. This, the convexity of the absolute value function and the observation $(X_1^{(n)}(t), \overline{X}_1^{(n)}(t)) \stackrel{d}{=} (X_2^{(n)}(t), \overline{X}_2^{(n)}(t)) \stackrel{d}{=} \cdots \stackrel{d}{=} (X_n^{(n)}(t), \overline{X}_n^{(n)}(t))$ allow to bound the expression in (3.13) from above by

$$
\begin{aligned}
(3.14) \quad & 2nC^2 \mathbb{E}\left[\mathbf{1}_{\{|X_1^{(n)}(t)| \geq |\overline{X}_1^{(n)}(t)|\}} e^{2C|X_1^{(n)}(t)|} (X_1^{(n)}(t) - \overline{X}_1^{(n)}(t))^2 \right] \\
& \quad + 2nC^2 \mathbb{E}\left[\mathbf{1}_{\{|\overline{X}_1^{(n)}(t)| > |X_1^{(n)}(t)|\}} e^{2C|\overline{X}_1^{(n)}(t)|} (X_1^{(n)}(t) - \overline{X}_1^{(n)}(t))^2 \right] \\
& \quad + 2C^2 \mathbb{E}[e^{2C|\overline{X}_1^{(n)}(t)|}].
\end{aligned}
$$

By dropping the indicator random variables, using the Cauchy–Schwarz inequality twice and invoking the estimate (2.8) and the $p = 4$ version of inequality (2.2) we conclude that the quantity in (3.14) is uniformly bounded in n and M. An analogous argument for the second expectation in (3.9) completes the proof of the lemma. $\square$

PROOF OF LEMMA 3.2.   With $F_{0,M}(x) := \int_0^x f_0(y) g_M(y) \,\mathrm{d}y$, we integrate by parts to rewrite the probability in (3.10) as

$$
\begin{aligned}
&\mathbb{P}\big(\big|\sqrt{n}\langle F_{0,M}, \varrho^{(n)}(t) - \varrho(t)\rangle\big| > \varepsilon\big) \\
(3.15) \qquad &\le \mathbb{P}\big(\big|\sqrt{n}\langle F_{0,M}, \varrho^{(n)}(t) - \overline{\varrho}^{(n)}(t)\rangle\big| > \varepsilon/2\big) \\
&\qquad + \mathbb{P}\big(\big|\sqrt{n}\langle F_{0,M}, \overline{\varrho}^{(n)}(t) - \varrho(t)\rangle\big| > \varepsilon/2\big)
\end{aligned}
$$

(note that the boundary terms in the integration by parts vanish thanks to $F_{0,M} \in \mathcal{E}_0$ and the estimate (2.8) in conjunction with Markov's inequality).

Now, we employ Markov's inequality, plug in the definitions of $\varrho^{(n)}(t)$, $\overline{\varrho}^{(n)}(t)$ and recall $(X_1^{(n)}(t), \overline{X}_1^{(n)}(t)) \overset{d}{=} (X_2^{(n)}(t), \overline{X}_2^{(n)}(t)) \overset{d}{=} \cdots \overset{d}{=} (X_n^{(n)}(t), \overline{X}_n^{(n)}(t))$ to control the first probability on the right-hand side of (3.15) by

$$
(3.16) \qquad \frac{2\sqrt{n}}{\varepsilon} \mathbb{E}\big[\big|F_{0,M}(X_1^{(n)}(t)) - F_{0,M}(\overline{X}_1^{(n)}(t))\big|\big].
$$

In view of the assumptions on f_0 and g_M, $M > 0$, we can find a constant $C < \infty$ independent of M such that $|f_0(x)| \le Ce^{C|x|}$, $x \in \mathbb{R}$ and $|g_M(x)| \le C\mathbf{1}_{\{|x|>M\}}$, $x \in \mathbb{R}$, $M > 0$. This and the convexity of the absolute value function show that the expression in (3.16) is not greater than

$$
\begin{aligned}
&\frac{2\sqrt{n}C^2}{\varepsilon} \mathbb{E}\big[e^{C|X_1^{(n)}(t)|} \mathbf{1}_{\{|X_1^{(n)}(t)|>M\}} \mathbf{1}_{\{|X_1^{(n)}(t)|\ge|\overline{X}_1^{(n)}(t)|\}} \big|X_1^{(n)}(t) - \overline{X}_1^{(n)}(t)\big|\big] \\
(3.17) \qquad &+ \frac{2\sqrt{n}C^2}{\varepsilon} \\
&\times \mathbb{E}\big[e^{C|\overline{X}_1^{(n)}(t)|} \mathbf{1}_{\{|\overline{X}_1^{(n)}(t)|>M\}} \mathbf{1}_{\{|\overline{X}_1^{(n)}(t)|>|X_1^{(n)}(t)|\}} \big|X_1^{(n)}(t) - \overline{X}_1^{(n)}(t)\big|\big].
\end{aligned}
$$

Leaving out the second indicator random variables from both expectations and applying Hölder's inequality twice we end up with

$$
\begin{aligned}
&\frac{2\sqrt{n}C^2}{\varepsilon} \mathbb{E}\big[e^{3C|X_1^{(n)}(t)|}\big]^{1/3} \mathbb{P}\big(|X_1^{(n)}(t)| > M\big)^{1/3} \mathbb{E}\big[|X_1^{(n)}(t) - \overline{X}_1^{(n)}(t)|^3\big]^{1/3} \\
(3.18) \qquad &+ \frac{2\sqrt{n}C^2}{\varepsilon} \mathbb{E}\big[e^{3C|\overline{X}_1^{(n)}(t)|}\big]^{1/3} \\
&\times \mathbb{P}\big(|\overline{X}_1^{(n)}(t)| > M\big)^{1/3} \mathbb{E}\big[|X_1^{(n)}(t) - \overline{X}_1^{(n)}(t)|^3\big]^{1/3},
\end{aligned}
$$

which tends to 0 when one takes the limits superior $n \to \infty$, $M \to \infty$ due to the estimates (2.8), (2.7) and the $p = 3$ version of inequality (2.2).

An appeal to Markov's inequality and the independence of $\overline{X}_1^{(n)}(t) \stackrel{d}{=} \overline{X}_2^{(n)}(t) \stackrel{d}{=} \cdots \stackrel{d}{=} \overline{X}_n^{(n)}(t) \stackrel{d}{=} \varrho(t)$ reveal that the second probability on the right-hand side of (3.15) is at most

$$(3.19) \qquad \frac{4}{\varepsilon^2} \mathbb{E}\big[\big(F_{0,M}\big(\overline{X}_1^{(n)}(t)\big) - \langle F_{0,M}, \varrho(t)\rangle\big)^2\big].$$

Moreover, by the definition of $F_{0,M}$, $M > 0$ and the assumptions on f_0 and g_M, $M > 0$ we have $F_{0,M}(x) = 0$, $x \in [-M, M]$, $M > 0$ and $|F_{0,M}(x)| \le Ce^{C|x|}$, $|x| > M$, $M > 0$, which allows to upper bound the latter expectation by

$$(3.20) \qquad \begin{aligned} & C^2 \mathbb{E}\big[\mathbf{1}_{\{|\overline{X}_1^{(n)}(t)|>M\}} e^{2C|\overline{X}_1^{(n)}(t)|}\big] \\ & \le C^2 \mathbb{P}\big(|\overline{X}_1^{(n)}(t)| > M\big)^{1/2} \mathbb{E}\big[e^{4C|\overline{X}_1^{(n)}(t)|}\big]^{1/2}. \end{aligned}$$

To complete the proof of the lemma, we pass to the limits superior $n \to \infty$, $M \to \infty$ relying on the estimates (2.7), (2.8) one more time. $\square$

CONVERGENCE OF $I_2^{(n)}(\cdot)$. We claim that, for all $t \ge 0$, it holds

$$(3.21) \qquad I_2^{(n)}(t) \stackrel{n\to\infty}{\longrightarrow} \big(\langle f_1, \varrho(t)\rangle, \ldots, \langle f_k, \varrho(t)\rangle\big)$$

in probability, which together with (3.1) and (3.5) yields Theorem 1.4. To obtain (3.21), we need to establish $\lim_{n\to\infty}\langle f_j, \widetilde{\varrho}^{(n)}(t)\rangle = \langle f_j, \varrho(t)\rangle$ in probability for every fixed $t \ge 0$ and $j \in \{1, 2, \ldots, k\}$. Consider the decomposition

$$(3.22) \qquad \begin{aligned} \langle f_j, \widetilde{\varrho}^{(n)}(t)\rangle = {}& \langle f_j, \varrho(t)\rangle + \xi^{(n)}(t)\langle f_j h_M, \varrho^{(n)}(t) - \varrho(t)\rangle \\ & + \xi^{(n)}(t)\langle f_j \widehat{h}_M, \varrho^{(n)}(t) - \varrho(t)\rangle, \end{aligned}$$

valid for any $M > 0$. We have $\lim_{n\to\infty} \xi^{(n)}(t)\langle f_j h_M, \varrho^{(n)}(t) - \varrho(t)\rangle = 0$ in probability due to $|\xi^{(n)}(t)| \le 1$ and Proposition 1.2 (note that $f_j h_M$ is continuous and bounded). Finally, $|\xi^{(n)}(t)| \le 1$, integration by parts (in which the boundary terms vanish thanks to the estimate (2.8) and Markov's inequality), the union bound and Lemma 3.2 give

$$\limsup_{M\to\infty} \limsup_{n\to\infty} \mathbb{P}\big(|\xi^{(n)}(t)\langle f_j \widehat{h}_M, \varrho^{(n)}(t) - \varrho(t)\rangle| > \varepsilon\big)$$

$$(3.23) \qquad \begin{aligned} & \le \limsup_{M\to\infty} \limsup_{n\to\infty} \mathbb{P}\bigg(\bigg|\int_{\mathbb{R}} f_j'(x)\widehat{h}_M(x)\big(F_{\varrho^{(n)}(t)}(x) - R(t,x)\big)\,\mathrm{d}x\bigg| > \varepsilon/2\bigg) \\ & \quad + \limsup_{M\to\infty} \limsup_{n\to\infty} \mathbb{P}\bigg(\bigg|\int_{\mathbb{R}} f_j(x)\widehat{h}_M'(x)\big(F_{\varrho^{(n)}(t)}(x) - R(t,x)\big)\,\mathrm{d}x\bigg| > \varepsilon/2\bigg) \\ & = 0 \end{aligned}$$

for all $\varepsilon > 0$, so that $\lim_{n\to\infty}\langle f_j, \widetilde{\varrho}^{(n)}(t)\rangle = \langle f_j, \varrho(t)\rangle$ in probability as desired. $\square$

4. Proof of Theorem 1.5.   With the simplified notation

$$(4.1) \qquad Z^{(n)}(\cdot) := \mathcal{J}_{J;\,f_1,\ldots,f_k}\big(\varrho^{(n)}(\cdot)\big) \quad \text{and} \quad Z(\cdot) := \mathcal{J}_{J;\,f_1,\ldots,f_k}\big(\varrho(\cdot)\big),$$

the main idea behind our proof of Theorem 1.5 lies in translating the convergence of $\sqrt{n}(\tau^{(n)} - \tau)$ to that of $\sqrt{n}(Z(\tau^{(n)}) - Z(\tau))$ by an application of the mean value theorem to $Z(\cdot)$ and in reducing the latter convergence to that of $\sqrt{n}(Z^{(n)}(\tau) - Z(\tau))$, for which Theorem 1.4 can be used. In this process, it is convenient to work with the truncated versions $\hat{\tau}^{(n)} := \tau^{(n)} \wedge (\tau + 1)$, $n \in \mathbb{N}$ of the hitting times $\tau^{(n)}$, $n \in \mathbb{N}$. The convergence in distribution of $\sqrt{n}(\tau^{(n)} - \tau)$ to a limit is then equivalent to the convergence in distribution of $\sqrt{n}(\hat{\tau}^{(n)} - \tau)$ to the same limit thanks to the following proposition, which is proved further below in this section.

PROPOSITION 4.1.   *In the setting of Theorem* 1.5, $\tau^{(n)} \overset{n \to \infty}{\longrightarrow} \tau$ *in probability.*

Our starting point for the proof of the convergence of $\sqrt{n}(\hat{\tau}^{(n)} - \tau)$ is the identity

$$(4.2) \qquad \mathbf{1}_{\{\tau^{(n)} \le \tau+1\}} Z^{(n)}\big(\hat{\tau}^{(n)}\big) = \mathbf{1}_{\{\tau^{(n)} \le \tau+1\}} Z(\tau).$$

The latter stems from the continuity of $Z^{(n)}(\cdot)$ and $Z(\cdot)$: for $Z^{(n)}(\cdot)$, it is a direct consequence of the definitions and, for $Z(\cdot)$, one can write $\langle f_1, \varrho(\cdot) \rangle, \ldots, \langle f_k, \varrho(\cdot) \rangle$ as $\mathbb{E}[f_1(\overline{X}_1^{(1)}(\cdot))], \ldots, \mathbb{E}[f_k(\overline{X}_1^{(1)}(\cdot))]$ and conclude by taking the expectation in Itô's formula and using Fubini's theorem (recall $f_1, \ldots, f_k \in \mathcal{E}_3 \subset \mathcal{E}_2$ and the estimate (2.8)). We observe in passing that, for the same reasons in conjunction with the dominated convergence theorem, $Z(\cdot)$ is actually continuously differentiable.

Next, we expand (4.2) into

$$
\begin{aligned}
(4.3) \qquad & \mathbf{1}_{\{\tau^{(n)} \le \tau+1\}} \sqrt{n}\big(Z(\hat{\tau}^{(n)}) - Z(\tau)\big) \\
& = -\mathbf{1}_{\{\tau^{(n)} \le \tau+1\}} \sqrt{n}\big(Z^{(n)}(\tau) - Z(\tau)\big) \\
& \quad + \mathbf{1}_{\{\tau^{(n)} \le \tau+1\}} \sqrt{n}\big(Z^{(n)}(\tau) - Z(\tau)\big) \\
& \quad - \mathbf{1}_{\{\tau^{(n)} \le \tau+1\}} \sqrt{n}\big(Z^{(n)}(\hat{\tau}^{(n)}) - Z(\hat{\tau}^{(n)})\big).
\end{aligned}
$$

In view of the continuous differentiability of $Z(\cdot)$, the mean value theorem and Proposition 4.1, the left-hand side of (4.3) converges in distribution as $n \to \infty$ if and only if $\sqrt{n}(\hat{\tau}^{(n)} - \tau)$ converges in distribution as $n \to \infty$, and the two limits differ by a factor of $Z'(\tau) \ne 0$ (cf. (1.12)). Concurrently, the first line on the right-hand side of (4.3) tends to $\sum_{j=1}^{k} \mathcal{J}_{J_{x_j};\,f_1,\ldots,f_k}(\varrho(\tau)) \int_{\mathbb{R}} f_j'(x) G(\tau, x)\, dx$ in distribution as $n \to \infty$ by Proposition 4.1 and Theorem 1.4.

To obtain Theorem 1.5, it now suffices to verify that the second line on the right-hand side of (4.3) converges to 0 in probability as $n \to \infty$. As a result of

(3.1), written in terms of $I_1^{(n)}(\cdot)$, $I_2^{(n)}(\cdot)$ from (3.2), (3.3), the desired convergence follows from

$$\sqrt{n}\big(Z^{(n)}(\tau) - Z(\tau)\big) - \sqrt{n}\big(Z^{(n)}(\widehat{\tau}^{(n)}) - Z(\widehat{\tau}^{(n)})\big)$$

$$= I_1^{(n)}(\tau)\nabla J\big(I_2^{(n)}(\tau)\big) - I_1^{(n)}(\widehat{\tau}^{(n)})\nabla J\big(I_2^{(n)}(\widehat{\tau}^{(n)})\big)$$

$$= I_1^{(n)}(\tau)\big(\nabla J\big(I_2^{(n)}(\tau)\big) - \nabla J\big(I_2^{(n)}(\widehat{\tau}^{(n)})\big)\big)$$

$$+ \big(I_1^{(n)}(\tau) - I_1^{(n)}(\widehat{\tau}^{(n)})\big)\nabla J\big(I_2^{(n)}(\widehat{\tau}^{(n)})\big),$$

(3.5), (3.21) and the next two lemmas.

LEMMA 4.2. *In the setting of Theorem* 1.5*,*

$$\nabla J\big(I_2^{(n)}(\widehat{\tau}^{(n)})\big) \stackrel{n\to\infty}{\Longrightarrow} \nabla J\big(\langle f_1, \varrho(\tau)\rangle, \ldots, \langle f_k, \varrho(\tau)\rangle\big) \tag{4.4}$$

in probability.

LEMMA 4.3. *In the setting of Theorem* 1.5*,* $I_1^{(n)}(\tau) - I_1^{(n)}(\widehat{\tau}^{(n)}) \stackrel{n\to\infty}{\Longrightarrow} 0$ *in probability.*

We complete the proof of Theorem 1.5 by establishing Proposition 4.1, Lemma 4.2 and Lemma 4.3.

PROOF OF PROPOSITION 4.1. Our proof of the proposition relies on the following lemma that extends the convergence result of Proposition 1.2 to test functions in $\mathcal{E}_0$.

LEMMA 4.4. *Let Assumption* 1.1 *be satisfied. Then, for all* $f_0 \in \mathcal{E}_0$, $T \geq 0$ *and* $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\Big(\sup_{t\in[0,T]} |\langle f_0, \varrho^{(n)}(t)\rangle - \langle f_0, \varrho(t)\rangle| > \varepsilon \Big) = 0. \tag{4.5}$$

PROOF. Recalling the auxiliary functions h_M, $M > 0$ and $\widehat{h}_M$, $M > 0$ from the beginning of Section 3, we know from Proposition 1.2 that, for any $M > 0$ and $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\Big(\sup_{t\in[0,T]} |\langle f_0 h_M, \varrho^{(n)}(t)\rangle - \langle f_0 h_M, \varrho(t)\rangle| > \varepsilon \Big) = 0. \tag{4.6}$$

Therefore, it is enough to check that

$$\limsup_{M\to\infty} \limsup_{n\to\infty} \mathbb{P}\Big(\sup_{t\in[0,T]} |\langle f_0 \widehat{h}_M, \varrho^{(n)}(t)\rangle| > \varepsilon \Big) = 0,$$

$$\limsup_{M\to\infty} \sup_{t\in[0,T]} |\langle f_0 \widehat{h}_M, \varrho(t)\rangle| = 0. \tag{4.7}$$

For the first assertion in (4.7), we use the definition of $\varrho^{(n)}(t)$, the observation $X_1^{(n)}(\cdot) \overset{d}{=} X_2^{(n)}(\cdot) \overset{d}{=} \cdots \overset{d}{=} X_n^{(n)}(\cdot)$, the estimate $|f_0(x)\widehat{h}_M(x)| \leq Ce^{C|x|}\mathbf{1}_{\{|x|>M\}}$, $x \in \mathbb{R}$ and the Cauchy–Schwarz inequality to deduce that

$$
\mathbb{E}\Big[\sup_{t\in[0,T]} |\langle f_0\widehat{h}_M, \varrho^{(n)}(t)\rangle| \Big]
$$

$$
= \mathbb{E}\Big[\sup_{t\in[0,T]} \Big| \frac{1}{n}\sum_{i=1}^{n} f_0(X_i^{(n)}(t))\widehat{h}_M(X_i^{(n)}(t)) \Big| \Big]
$$

(4.8)
$$
\leq \mathbb{E}\Big[\sup_{t\in[0,T]} |f_0(X_1^{(n)}(t))\widehat{h}_M(X_1^{(n)}(t))| \Big]
$$

$$
\leq C\mathbb{E}\Big[\sup_{t\in[0,T]} \big(e^{C|X_1^{(n)}(t)|}\mathbf{1}_{\{|X_1^{(n)}(t)|>M\}} \big) \Big]
$$

$$
\leq C\mathbb{E}\Big[\sup_{t\in[0,T]} e^{2C|X_1^{(n)}(t)|} \Big]^{1/2} \mathbb{E}\Big[\sup_{t\in[0,T]} \mathbf{1}_{\{|X_1^{(n)}(t)|>M\}} \Big]^{1/2}.
$$

Since $\sup_{t\in[0,T]} \mathbf{1}_{\{|X_1^{(n)}(t)|>M\}} = \mathbf{1}_{\{\sup_{t\in[0,T]} |X_1^{(n)}(t)|>M\}}$, the first assertion in (4.7) now follows from Markov's inequality and the estimates (2.8), (2.7).

For the second assertion in (4.7), we recall that $\overline{X}_1^{(n)}(t) \overset{d}{=} \varrho(t)$, $t \geq 0$, allowing us to bound $\sup_{t\in[0,T]} |\langle f_0\widehat{h}_M, \varrho(t)\rangle|$ by

$$
\mathbb{E}\Big[\sup_{t\in[0,T]} |f_0(\overline{X}_1^{(n)}(t))\widehat{h}_M(\overline{X}_1^{(n)}(t))| \Big]
$$

(4.9)
$$
\leq C\mathbb{E}\Big[\sup_{t\in[0,T]} e^{2C|\overline{X}_1^{(n)}(t)|} \Big]^{1/2} \mathbb{P}\Big(\sup_{t\in[0,T]} |\overline{X}_1^{(n)}(t)| > M \Big)^{1/2}
$$

via the procedure in the last paragraph. The estimates (2.8), (2.7) yield the result. $\qquad\square$

Back to the proof of Proposition 4.1, given Lemma 4.4, the uniform continuity of the function J on compact neighborhoods of the set

$$
\{(\langle f_1, \varrho(t)\rangle, \ldots, \langle f_k, \varrho(t)\rangle) : t \in [0, \tau+1]\} \subset \mathbb{R}^k
$$

implies

(4.10)
$$
\lim_{n\to\infty} \mathbb{P}\Big(\sup_{t\in[0,\tau+1]} |Z^{(n)}(t) - Z(t)| > \varepsilon \Big) = 0, \quad \varepsilon > 0.
$$

Since $Z'(\tau) \neq 0$ (cf. (1.12)) and $Z'(\cdot)$ is continuous, there exist $[0,1] \ni \upsilon_m \downarrow 0$ such that $Z(\tau + \upsilon_m) > Z(\tau)$ for all $m \in \mathbb{N}$ if $Z'(\tau) > 0$, or $Z(\tau + \upsilon_m) < Z(\tau)$ for all $m \in \mathbb{N}$ if $Z'(\tau) < 0$. Applying (4.10) with $\varepsilon := |Z(\tau + \upsilon_m) - Z(\tau)|/2 = |Z(\tau + \upsilon_m) - a|/2$ consecutively, we find that

(4.11)
$$
\lim_{n\to\infty} \mathbb{P}\big(\tau^{(n)} > \tau + \upsilon_m \big) = 0
$$

for all m. At the same time, for all $\upsilon > 0$, we have by (4.10):

$$
\text{(4.12)} \quad \begin{aligned}
&\lim_{n \to \infty} \mathbb{P}(\tau^{(n)} \leq \tau - \upsilon) \\
&\leq \lim_{n \to \infty} \mathbb{P}\Big(\sup_{t \in [0, \tau - \upsilon]} |Z^{(n)}(t) - Z(t)| \geq \min_{t \in [0, \tau - \upsilon]} |a - Z(t)| \Big) = 0.
\end{aligned}
$$

This completes the proof of Proposition 4.1. $\quad \square$

PROOF OF LEMMA 4.2. In view of the continuity of ∇J, it suffices to show that

$$
\text{(4.13)} \quad \langle f_j, \widetilde{\varrho}^{(n)}(\widehat{\tau}^{(n)})\rangle \overset{n \to \infty}{\longrightarrow} \langle f_j, \varrho(\tau)\rangle, \quad j = 1, 2, \ldots, k
$$

in probability. Since $\widetilde{\varrho}^{(n)}(\widehat{\tau}^{(n)})$ is a convex combination of $\varrho^{(n)}(\widehat{\tau}^{(n)})$ and $\varrho(\widehat{\tau}^{(n)})$, we may swap $\widetilde{\varrho}^{(n)}(\widehat{\tau}^{(n)})$ for $\varrho(\widehat{\tau}^{(n)})$ on the left-hand side of (4.13) by Lemma 4.4 with $T := \tau + 1$. Then Proposition 4.1 and the continuity of $\langle f_j, \varrho(\cdot)\rangle$ give the lemma. $\quad \square$

PROOF OF LEMMA 4.3. We need to verify that

$$
\text{(4.14)} \quad \begin{aligned}
&\sqrt{n}\langle f_j, \varrho^{(n)}(\tau) - \varrho^{(n)}(\widehat{\tau}^{(n)})\rangle - \sqrt{n}\langle f_j, \varrho(\tau) - \varrho(\widehat{\tau}^{(n)})\rangle \overset{n \to \infty}{\longrightarrow} 0, \\
&j = 1, 2, \ldots, k
\end{aligned}
$$

in probability. For a fixed $j \in \{1, 2, \ldots, k\}$, we start by establishing the corresponding convergence under the assumption that $f_j \in C_c^3(\mathbb{R}) \subset \mathcal{E}_3$.

Step 1: convergence (4.14) *for* $f_j \in C_c^3(\mathbb{R})$. Inserting the definition of $\varrho^{(n)}(\cdot)$ and applying Itô's formula, we find for the first term in (4.14):

$$
\text{(4.15)} \quad \begin{aligned}
&\sqrt{n}\langle f_j, \varrho^{(n)}(\tau) - \varrho^{(n)}(\widehat{\tau}^{(n)})\rangle \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big(f_j(X_i^{(n)}(\tau)) - f_j(X_i^{(n)}(\widehat{\tau}^{(n)}))\big) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{\widehat{\tau}^{(n)}}^{\tau} b\big(F_{\varrho^{(n)}(t)}(X_i^{(n)}(t))\big) f_j'(X_i^{(n)}(t)) \\
&\quad + \frac{\sigma(F_{\varrho^{(n)}(t)}(X_i^{(n)}(t)))^2}{2} f_j''(X_i^{(n)}(t)) \, dt \\
&\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{\widehat{\tau}^{(n)}}^{\tau} \sigma\big(F_{\varrho^{(n)}(t)}(X_i^{(n)}(t))\big) f_j'(X_i^{(n)}(t)) \, dB_i^{(n)}(t).
\end{aligned}
$$

To simplify the second line in (4.15), we introduce the discrete antiderivatives

$$(\mathcal{I}_n b)(r) := \frac{1}{n} \sum_{i=1}^{n} b(i/n) \mathbf{1}_{\{r \geq i/n\}},$$

(4.16)

$$\left(\mathcal{I}_n \frac{\sigma^2}{2}\right)(r) := \frac{1}{n} \sum_{i=1}^{n} \frac{\sigma(i/n)^2}{2} \mathbf{1}_{\{r \geq i/n\}}, \quad r \in [0, 1].$$

Since the order statistics $X_{(1)}^{(n)}(t) \leq X_{(2)}^{(n)}(t) \leq \cdots \leq X_{(n)}^{(n)}(t)$ are almost surely distinct for Lebesgue almost every $t \geq 0$ by [23], theorem on page 439, for the function $x \mapsto \sum_{1 \leq i_1 < i_2 \leq n} \mathbf{1}_{\{x_{i_1} = x_{i_2}\}}$, we can now use summation by parts, the piecewise constant nature of $(\mathcal{I}_n b)(F_{\varrho^{(n)}(t)}(\cdot))$, and the convention $X_{(n+1)}^{(n)}(t) = \infty$ to compute

$$\frac{1}{n} \sum_{i=1}^{n} b\big(F_{\varrho^{(n)}(t)}(X_i^{(n)}(t))\big) f_j'(X_i^{(n)}(t))$$

$$= \frac{1}{n} \sum_{i=1}^{n} b(i/n) f_j'(X_{(i)}^{(n)}(t))$$

(4.17)

$$= \sum_{i=1}^{n} \big((\mathcal{I}_n b)(i/n) - (\mathcal{I}_n b)((i-1)/n)\big) f_j'(X_{(i)}^{(n)}(t))$$

$$= \sum_{i=1}^{n} (\mathcal{I}_n b)(i/n) \big(f_j'(X_{(i)}^{(n)}(t)) - f_j'(X_{(i+1)}^{(n)}(t))\big)$$

$$= - \int_{\mathbb{R}} (\mathcal{I}_n b)\big(F_{\varrho^{(n)}(t)}(x)\big) f_j''(x) \, dx.$$

Similarly, we see that

$$\frac{1}{n} \sum_{i=1}^{n} \frac{\sigma(F_{\varrho^{(n)}(t)}(X_i^{(n)}(t)))^2}{2} f_j''(X_i^{(n)}(t))$$

(4.18)

$$= - \int_{\mathbb{R}} \left(\mathcal{I}_n \frac{\sigma^2}{2}\right)\big(F_{\varrho^{(n)}(t)}(x)\big) f_j'''(x) \, dx.$$

Consequently, we arrive at

$$\sqrt{n} \langle f_j, \varrho^{(n)}(\tau) - \varrho^{(n)}(\widehat{\tau}^{(n)}) \rangle$$

$$= -\sqrt{n} \int_{\widehat{\tau}^{(n)}}^{\tau} \int_{\mathbb{R}} (\mathcal{I}_n b)\big(F_{\varrho^{(n)}(t)}(x)\big) f_j''(x)$$

(4.19)

$$+ \left(\mathcal{I}_n \frac{\sigma^2}{2}\right)\big(F_{\varrho^{(n)}(t)}(x)\big) f_j'''(x) \, dx \, dt$$

$$+ N^{(n)}(\tau) - N^{(n)}(\widehat{\tau}^{(n)}),$$

where

$$(4.20) \quad N^{(n)}(t) := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^t \sigma\big(F_{\varrho^{(n)}(s)}(X_i^{(n)}(s))\big) f_j'(X_i^{(n)}(s)) \, dB_i^{(n)}(s), \quad t \geq 0.$$

On the other hand, integration by parts and the notion of a generalized solution for the PDE (1.5) (see [14], Definition 3) imply that

$$(4.21) \quad \begin{aligned} \sqrt{n} \langle f_j, \varrho(\tau) - \varrho(\widehat{\tau}^{(n)}) \rangle = &-\sqrt{n} \int_{\widehat{\tau}^{(n)}}^{\tau} \int_{\mathbb{R}} B\big(R(t,x)\big) f_j''(x) \\ &+ \Sigma\big(R(t,x)\big) f_j'''(x) \, dx \, dt, \end{aligned}$$

which can be combined with (4.19) to

$$(4.22) \quad \begin{aligned} \sqrt{n} \langle f_j, &\varrho^{(n)}(\tau) - \varrho^{(n)}(\widehat{\tau}^{(n)}) \rangle - \sqrt{n} \langle f_j, \varrho(\tau) - \varrho(\widehat{\tau}^{(n)}) \rangle \\ = &-\sqrt{n} \int_{\widehat{\tau}^{(n)}}^{\tau} \int_{\mathbb{R}} \big((\mathcal{I}_n b)(F_{\varrho^{(n)}(t)}(x)) - B\big(R(t,x)\big)\big) f_j''(x) \\ &+ \left(\left(\mathcal{I}_n \frac{\sigma^2}{2}\right)(F_{\varrho^{(n)}(t)}(x)) - \Sigma\big(R(t,x)\big)\right) f_j'''(x) \, dx \, dt \\ &+ N^{(n)}(\tau) - N^{(n)}(\widehat{\tau}^{(n)}). \end{aligned}$$

To prove that the right-hand side of (4.22) converges to 0 in probability we note that the Lipschitz property of b, $\frac{\sigma^2}{2}$ (cf. Assumption 1.1(b)) yields

$$(4.23) \quad \begin{aligned} \lim_{n\to\infty} \sqrt{n} \sup_{r\in[0,1]} &\big|(\mathcal{I}_n b)(r) - B(r)\big| \\ &= \lim_{n\to\infty} \sqrt{n} \sup_{r\in[0,1]} \left|\left(\mathcal{I}_n \frac{\sigma^2}{2}\right)(r) - \Sigma(r)\right| = 0. \end{aligned}$$

Since, in addition, B, Σ are Lipschitz (cf. Assumption 1.1(b)) and f_j'', f_j''' are bounded, it suffices to obtain the limits in probability

$$(4.24) \quad \begin{aligned} \lim_{n\to\infty} \sqrt{n} \int_{\widehat{\tau}^{(n)}}^{\tau} \int_{\mathbb{R}} \big|F_{\varrho^{(n)}(t)}(x) - R(t,x)\big| \, dx \, dt = 0 \quad \text{and} \\ \lim_{n\to\infty} \big(N^{(n)}(\tau) - N^{(n)}(\widehat{\tau}^{(n)})\big) = 0. \end{aligned}$$

For the first convergence in (4.24), we recall the representation of the W_1-distance in (2.9) and apply the triangle inequality for the latter together with

Markov's inequality and Fubini's theorem to find, for all $\varepsilon, \varepsilon' > 0$,

$$\mathbb{P}\left(\left|\sqrt{n}\int_{\widehat{\tau}^{(n)}}^{\tau} W_1(\varrho^{(n)}(t), \varrho(t))\,\mathrm{d}t\right| > \varepsilon\right)$$

(4.25)
$$\leq \mathbb{P}(|\widehat{\tau}^{(n)} - \tau| > \varepsilon') + \mathbb{P}\left(\sqrt{n}\int_{\tau-\varepsilon'}^{\tau+\varepsilon'} W_1(\varrho^{(n)}(t), \varrho(t))\,\mathrm{d}t > \varepsilon\right)$$

$$\leq \mathbb{P}(|\widehat{\tau}^{(n)} - \tau| > \varepsilon') + \frac{\sqrt{n}}{\varepsilon}\int_{\tau-\varepsilon'}^{\tau+\varepsilon'} \mathbb{E}[W_1(\varrho^{(n)}(t), \overline{\varrho}^{(n)}(t))]$$

$$+ \mathbb{E}[W_1(\overline{\varrho}^{(n)}(t), \varrho(t))]\,\mathrm{d}t.$$

In view of Propositions 4.1, 2.1 and 2.3, this estimate tends to 0 for all $\varepsilon > 0$ when we take $n \to \infty$ and then $\varepsilon' \downarrow 0$.

For the second convergence in (4.24), we compute the quadratic variation process

$$(4.26) \qquad [N^{(n)}](t) = \int_0^t \frac{1}{n}\sum_{i=1}^n \sigma\big(F_{\varrho^{(n)}(s)}(X_i^{(n)}(s))\big)^2 f_j'(X_i^{(n)}(s))^2\,\mathrm{d}s, \quad t \geq 0,$$

bound the resulting integrand by a constant $C < \infty$, and use the martingale representation theorem (see, e.g., [21], Chapter 3, Theorem 4.6 and Problem 4.7) to conclude

$$(4.27) \quad \mathbb{P}(|N(\tau) - N(\widehat{\tau}^{(n)})| > \varepsilon) \leq \mathbb{P}(|\widehat{\tau}^{(n)} - \tau| > \varepsilon') + \mathbb{P}\Big(\sup_{t \in [0, C\varepsilon']} |B_1^{(1)}(t)| > \varepsilon\Big).$$

Thanks to Proposition 4.1, it is now enough to send $n \to \infty$ followed by $\varepsilon' \downarrow 0$.

Step 2: convergence (4.14) *for general* $f_j \in \mathcal{E}_3$. With the functions $h_M, M > 0$ and $\widehat{h}_M, M > 0$ introduced at the beginning of Section 3, we decompose the left-hand side of (4.14) into

(4.28)
$$\sqrt{n}\langle f_j h_M, \varrho^{(n)}(\tau) - \varrho^{(n)}(\widehat{\tau}^{(n)})\rangle - \sqrt{n}\langle f_j h_M, \varrho(\tau) - \varrho(\widehat{\tau}^{(n)})\rangle$$
$$+ \sqrt{n}\langle f_j \widehat{h}_M, \varrho^{(n)}(\tau) - \varrho^{(n)}(\widehat{\tau}^{(n)})\rangle - \sqrt{n}\langle f_j \widehat{h}_M, \varrho(\tau) - \varrho(\widehat{\tau}^{(n)})\rangle.$$

The first line in (4.28) converges to 0 in probability as $n \to \infty$ by Step 1, so we focus on the second line in (4.28). To move from $\varrho^{(n)}(\cdot)$ to $\overline{\varrho}^{(n)}(\cdot)$ therein, we will prove that, for all $\varepsilon > 0$,

(4.29)
$$\limsup_{M \to \infty}\limsup_{n \to \infty} \mathbb{P}(|\sqrt{n}\langle f_j \widehat{h}_M, \varrho^{(n)}(\tau) - \varrho^{(n)}(\widehat{\tau}^{(n)})$$
$$- \overline{\varrho}^{(n)}(\tau) + \overline{\varrho}^{(n)}(\widehat{\tau}^{(n)})\rangle| > \varepsilon) = 0.$$

We recall that $\widehat{\tau}^{(n)} \in [0, \tau + 1]$, insert the definitions of $\varrho^{(n)}$, $\overline{\varrho}^{(n)}$ and exploit Markov's inequality to bound the probability in (4.29) by

(4.30)
$$\frac{2\sqrt{n}}{\varepsilon} \mathbb{E}\left[\sup_{t \in [0, \tau+1]} \left| \frac{1}{n} \sum_{i=1}^{n} ((f_j \widehat{h}_M)(X_i^{(n)}(t)) - (f_j \widehat{h}_M)(\overline{X}_i^{(n)}(t))) \right| \right]$$
$$\leq \frac{2\sqrt{n}}{\varepsilon} \mathbb{E}\left[\sup_{t \in [0, \tau+1]} |(f_j \widehat{h}_M)(X_1^{(n)}(t)) - (f_j \widehat{h}_M)(\overline{X}_1^{(n)}(t))| \right],$$

where we have used $(X_1^{(n)}(\cdot), \overline{X}_1^{(n)}(\cdot)) \overset{d}{=} (X_2^{(n)}(\cdot), \overline{X}_2^{(n)}(\cdot)) \overset{d}{=} \cdots \overset{d}{=} (X_n^{(n)}(\cdot),$ $\overline{X}_n^{(n)}(\cdot))$. Due to the mean value theorem for $f_j \widehat{h}_M$, the inequality $|(f_j \widehat{h}_M)'(x)| \leq C e^{C|x|} \mathbf{1}_{\{|x|>M\}}$, $x \in \mathbb{R}$, the convexity of the absolute value function and Hölder's inequality the right-hand side of (4.30) is less or equal to

(4.31)
$$\frac{2\sqrt{n}}{\varepsilon} \mathbb{E}\left[\sup_{t \in [0, \tau+1]} |X_1^{(n)}(t) - \overline{X}_1^{(n)}(t)|^3 \right]^{1/3}$$
$$\cdot \left(C \mathbb{E}\left[\sup_{t \in [0, \tau+1]} e^{3C|X_1^{(n)}(t)|} \right]^{1/3} \mathbb{P}\left(\sup_{t \in [0, \tau+1]} |X_1^{(n)}(t)| > M \right)^{1/3} \right.$$
$$\left. + C \mathbb{E}\left[\sup_{t \in [0, \tau+1]} e^{3C|\overline{X}_1^{(n)}(t)|} \right]^{1/3} \mathbb{P}\left(\sup_{t \in [0, \tau+1]} |\overline{X}_1^{(n)}(t)| > M \right)^{1/3} \right).$$

At this point, (4.29) becomes a consequence of the inequality (2.2) with $p = 3$ and the estimates (2.8), (2.7).

With $\mathcal{A}_t := b(R(t, \cdot)) \frac{\mathrm{d}}{\mathrm{d}x} + \frac{\sigma(R(t, \cdot))^2}{2} \frac{\mathrm{d}^2}{\mathrm{d}x^2}$, $t \geq 0$, we may now replace $\varrho^{(n)}(\cdot)$ by $\overline{\varrho}^{(n)}(\cdot)$ in the second line of (4.28) and deduce by means of Itô's formula that

(4.32)
$$\sqrt{n} \langle f_j \widehat{h}_M, \overline{\varrho}^{(n)}(\tau) - \overline{\varrho}^{(n)}(\widehat{\tau}^{(n)}) \rangle$$
$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} ((f_j \widehat{h}_M)(\overline{X}_i^{(n)}(\tau)) - (f_j \widehat{h}_M)(\overline{X}_i^{(n)}(\widehat{\tau}^{(n)})))$$
$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{\widehat{\tau}^{(n)}}^{\tau} (\mathcal{A}_t(f_j \widehat{h}_M))(\overline{X}_i^{(n)}(t)) \, \mathrm{d}t + \overline{N}^{(n)}(\tau) - \overline{N}^{(n)}(\widehat{\tau}^{(n)}),$$

where

(4.33)
$$\overline{N}^{(n)}(t) := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{0}^{t} \sigma(R(s, \overline{X}_i^{(n)}(s)))(f_j \widehat{h}_M)'(\overline{X}_i^{(n)}(s)) \, \mathrm{d}B_i^{(n)}(s),$$
$$t \geq 0.$$

Next, we take the expectation in Itô's formula for $(f_j \widehat{h}_M)(\overline{X}_1^{(1)}(\tau)) - (f_j \widehat{h}_M) \times (\overline{X}_1^{(1)}(t))$, $t \geq 0$ relying on $\overline{X}_1^{(1)}(t) \overset{d}{=} \varrho(t)$, $t \geq 0$ and employ Fubini's theorem

(note $f_j \widehat{h}_M \in \mathcal{E}_2$ and the estimate (2.8)), followed by an evaluation at $t = \widehat{\tau}^{(n)}$ to get

$$(4.34) \qquad \langle f_j \widehat{h}_M, \varrho(\tau) - \varrho(\widehat{\tau}^{(n)}) \rangle = \int_{\widehat{\tau}^{(n)}}^{\tau} \mathbb{E}\big[(\mathcal{A}_t(f_j \widehat{h}_M))(\overline{X}_1^{(1)}(t))\big] \, dt.$$

We proceed using the union bound, Markov's inequality and $\widehat{\tau}^{(n)} \in [0, \tau + 1]$:

$$\mathbb{P}\big(\big|\sqrt{n}\langle f_j \widehat{h}_M, \overline{\varrho}^{(n)}(\tau) - \overline{\varrho}^{(n)}(\widehat{\tau}^{(n)})\rangle - \sqrt{n}\langle f_j \widehat{h}_M, \varrho(\tau) - \varrho(\widehat{\tau}^{(n)})\rangle\big| > \varepsilon\big)$$

$$\leq \frac{2}{\varepsilon} \mathbb{E}\Bigg[\int_0^{\tau+1} \Bigg| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} ((\mathcal{A}_t(f_j \widehat{h}_M))(\overline{X}_i^{(n)}(t))$$

$$(4.35) \qquad\qquad - \mathbb{E}[(\mathcal{A}_t(f_j \widehat{h}_M))(\overline{X}_i^{(n)}(t))])\Bigg| \, dt\Bigg]$$

$$+ \mathbb{P}\big(\big|\overline{N}^{(n)}(\tau) - \overline{N}^{(n)}(\widehat{\tau}^{(n)})\big| > \varepsilon/2\big)$$

$$\leq \frac{2}{\varepsilon} \int_0^{\tau+1} \mathbb{SD}((\mathcal{A}_t(f_j \widehat{h}_M))(\overline{X}_1^{(1)}(t))) \, dt$$

$$+ \mathbb{P}\big(\big|\overline{N}^{(n)}(\tau) - \overline{N}^{(n)}(\widehat{\tau}^{(n)})\big| > \varepsilon/2\big),$$

where we have applied Fubini's theorem, Jensen's inequality and the independence of $\overline{X}_1^{(n)}(t) \overset{d}{=} \overline{X}_2^{(n)}(t) \overset{d}{=} \cdots \overset{d}{=} \overline{X}_n^{(n)}(t) \overset{d}{=} \varrho(t)$ and have written $\mathbb{SD}$ for the standard deviation operator. Since the standard deviation of a random variable does not exceed its L^2-norm, the boundedness of b, σ (cf. Assumption 1.1(b)), $f_j \in \mathcal{E}_2$ and the properties of $\widehat{h}_M$ imply an estimate of the form

$$(4.36) \qquad \mathbb{SD}((\mathcal{A}_t(f_j \widehat{h}_M))(\overline{X}_1^{(1)}(t))) \leq \mathbb{E}\big[C e^{C|\overline{X}_1^{(1)}(t)|} \mathbf{1}_{\{|\overline{X}_1^{(1)}(t)| > M\}}\big]^{1/2}.$$

Moreover, its right-hand side tends to 0 as $M \to \infty$ uniformly in $t \in [0, \tau]$ by the Cauchy–Schwarz inequality and the estimates (2.8), (2.7).

To complete the proof, we need to analyze the last probability in (4.35). For this purpose, we compute

$$(4.37) \quad [\overline{N}^{(n)}](t) = \frac{1}{n} \sum_{i=1}^{n} \int_0^t \sigma\big(R(s, \overline{X}_i^{(n)}(s))\big)^2 (f_j \widehat{h}_M)'(\overline{X}_i^{(n)}(s))^2 \, ds, \quad t \geq 0.$$

Thus, the martingale representation theorem (see, e.g., [21], Chapter 3, Theorem 4.6 and Problem 4.7) and the union bound give, for all $\varepsilon' > 0$,

$$\mathbb{P}\big(\big|\overline{N}^{(n)}(\tau) - \overline{N}^{(n)}(\widehat{\tau}^{(n)})\big| > \varepsilon/2\big)$$

$$(4.38) \qquad \leq \mathbb{P}\big(\big|[\overline{N}^{(n)}](\tau) - [\overline{N}^{(n)}](\widehat{\tau}^{(n)})\big| > \varepsilon'\big) + \mathbb{P}\Big(\sup_{t \in [0, \varepsilon']} |B_1^{(1)}(t)| > \varepsilon/2\Big).$$

Due to $\widehat{\tau}^{(n)} \in [0, \tau+1]$, Markov's inequality, $\overline{X}_1^{(n)}(\cdot) \stackrel{d}{=} \overline{X}_2^{(n)}(\cdot) \stackrel{d}{=} \cdots \stackrel{d}{=} \overline{X}_n^{(n)}(\cdot) \stackrel{d}{=} \overline{X}_1^{(1)}(\cdot)$, Fubini's theorem, the boundedness of σ (cf. Assumption 1.1(b)), $f_j \in \mathcal{E}_1$ and the properties of $\widehat{h}_M$ we have an estimate of the type

$$
(4.39) \quad
\begin{aligned}
&\mathbb{P}\big(|[\overline{N}^{(n)}](\tau) - [\overline{N}^{(n)}](\widehat{\tau}^{(n)})| > \varepsilon'\big) \\
&\qquad \le \frac{1}{\varepsilon'} \int_0^{\tau+1} \mathbb{E}\big[C e^{C|\overline{X}_1^{(1)}(t)|} \mathbf{1}_{\{|\overline{X}_1^{(1)}(t)|>M\}}\big] \mathrm{d}t.
\end{aligned}
$$

The latter converges to 0 as $M \to \infty$ thanks to the Cauchy–Schwarz inequality and the estimates (2.8), (2.7). It remains to observe that the second probability on the right-hand side of (4.38) vanishes as $\varepsilon' \downarrow 0$. $\square$

5. Applications in stochastic portfolio theory.

5.1. *Dynamics of the market diversity.* Consider a stock market with n companies, as described by the market weight processes $\mu_1(\cdot), \mu_2(\cdot), \ldots, \mu_n(\cdot)$, that is, the fractions of the total market capital invested in the different companies at any given time. In this context, a concept that has attracted much interest, both for scientific reasons and its importance in investment decisions, is the *market diversity*. Informally speaking, a market is thought of as diverse when one can be certain that no single company will end up with the vast majority of the market capital. In [10], Fernholz has proposed to formalize the notion of diversity as follows.

DEFINITION 5.1 ([10], Definition 2.2.1). *A market is called* diverse *if for some $\varepsilon > 0$ it holds $\max_{1 \le i \le n} \mu_i(t) \le 1 - \varepsilon$ for all $t \ge 0$ almost surely. A market is referred to as* weakly diverse *on a finite time interval $[0, T]$ if for some $\varepsilon > 0$ one has*

$$
(5.1) \quad \frac{1}{T} \int_0^T \max_{1 \le i \le n} \mu_i(t) \, \mathrm{d}t \le 1 - \varepsilon
$$

almost surely.

Subsequently, it is noticed in [10] that the vector of the market weight processes $\mu(\cdot) := (\mu_1(\cdot), \mu_2(\cdot), \ldots, \mu_n(\cdot))$ takes values in the closed unit simplex

$$
(5.2) \quad \overline{\Delta}^n := \left\{ x \in [0,1]^n : \sum_{i=1}^n x_i = 1 \right\},
$$

whereas the diversity condition $\max_{1 \le i \le n} \mu_i(t) \le 1 - \varepsilon$ is violated when $\mu(\cdot)$ enters the corresponding open neighborhoods of the vertices of $\overline{\Delta}^n$. Hence, it is natural to use a symmetric concave function on $\overline{\Delta}^n$, which necessarily attains its minimum at the vertices, to quantify the diversity of a market (or the lack thereof). The main examples of such functions discussed in [10] are:

(i) the entropy function $H(t) = -\sum_{i=1}^{n} \mu_i(t) \log \mu_i(t)$, $t \geq 0$,
(ii) the ℓ^p-norms $D_p(t) = (\sum_{i=1}^{n} \mu_i(t)^p)^{1/p}$, $t \geq 0$ for $p \in (0, 1)$,
(iii) and the geometric mean $S(t) = (\prod_{i=1}^{n} \mu_i(t))^{1/n}$, $t \geq 0$.

In particular, the entropy function and the ℓ^p-norms for $p \in (0, 1)$ can be employed to test if a market is diverse in the sense of Definition 5.1 (cf. [10], Proposition 2.3.2).

PROPOSITION 5.2. *A market is diverse if and only if for some $\varepsilon' > 0$ it holds $H(t) \geq \varepsilon'$ for all $t \geq 0$ almost surely or, equivalently, for some $p \in (0, 1)$ and $\varepsilon'' > 0$ one has $D_p(t) \geq 1 + \varepsilon''$ for all $t \geq 0$ almost surely.*

Our Theorem 1.4 can be utilized to capture the dynamics of the entropy $H(\cdot)$, the ℓ^p-norms $D_p(\cdot)$, $p \in (0, 1)$ and the geometric mean $S(\cdot)$ in rank-based models with a large number n of companies. In that setting, the market weight processes are defined in terms of the solution to (1.3) by

$$(5.3) \qquad \mu_i(\cdot) = \frac{e^{X_i^{(n)}(\cdot)}}{e^{X_1^{(n)}(\cdot)} + e^{X_2^{(n)}(\cdot)} + \cdots + e^{X_n^{(n)}(\cdot)}}, \quad i = 1, 2, \ldots, n$$

and give rise to the associated entropy, ℓ^p-norm and geometric mean processes via the items (i), (ii), (iii) above.

COROLLARY 5.3. *Under Assumption 1.1, the following convergences hold in the finite-dimensional distribution sense:*

(a) *for the entropy process $H(\cdot)$,*

$$\sqrt{n}\left(H(\cdot) - \log n - \log\langle e^x, \varrho(\cdot)\rangle + \frac{\langle x e^x, \varrho(\cdot)\rangle}{\langle e^x, \varrho(\cdot)\rangle} \right)$$

$$(5.4) \qquad \xrightarrow{n \to \infty} -\left(\frac{1}{\langle e^x, \varrho(\cdot)\rangle} + \frac{\langle x e^x, \varrho(\cdot)\rangle}{\langle e^x, \varrho(\cdot)\rangle^2} \right) \int_{\mathbb{R}} e^x G(\cdot, x) \, dx$$

$$- \frac{1}{\langle e^x, \varrho(\cdot)\rangle} \int_{\mathbb{R}} (e^x + x e^x) G(\cdot, x) \, dx,$$

(b) *for an ℓ^p-norm process $D_p(\cdot)$ with $p \in (0, 1)$,*

$$\sqrt{n}\left(n^{\frac{p-1}{p}} D_p(\cdot) - \frac{\langle e^{px}, \varrho(\cdot)\rangle^{1/p}}{\langle e^x, \varrho(\cdot)\rangle} \right)$$

$$(5.5) \qquad \xrightarrow{n \to \infty} -\frac{\langle e^{px}, \varrho(\cdot)\rangle^{1/p-1}}{\langle e^x, \varrho(\cdot)\rangle} \int_{\mathbb{R}} e^{px} G(\cdot, x) \, dx$$

$$+ \frac{\langle e^{px}, \varrho(\cdot)\rangle^{1/p}}{\langle e^x, \varrho(\cdot)\rangle^2} \int_{\mathbb{R}} e^x G(\cdot, x) \, dx,$$

(c) *for the the geometric mean process $S(\cdot)$,*

$$\sqrt{n}\left(nS(\cdot) - \frac{e^{\langle x, \varrho(\cdot)\rangle}}{\langle e^x, \varrho(\cdot)\rangle}\right) \overset{n\to\infty}{\longrightarrow} -\frac{e^{\langle x, \varrho(\cdot)\rangle}}{\langle e^x, \varrho(\cdot)\rangle}\int_{\mathbb{R}} G(\cdot, x)\, dx$$

(5.6)

$$+ \frac{e^{\langle x, \varrho(\cdot)\rangle}}{\langle e^x, \varrho(\cdot)\rangle^2}\int_{\mathbb{R}} e^x G(\cdot, x)\, dx.$$

PROOF. The corollary is a direct consequence of Theorem 1.4. For the sake of completeness, we write out the functions J, $f_1, \ldots, f_k$ in each of the three cases.

(a) For the normalized entropy process $H(\cdot) - \log n$, take

$$J : (0, \infty) \times \mathbb{R} \to \mathbb{R}, \qquad (x_1, x_2) \mapsto \log x_1 - \frac{x_2}{x_1},$$

(5.7)

$$f_1(x) = e^x, \qquad f_2(x) = x e^x.$$

(b) For every normalized ℓ^p-norm process $n^{\frac{p-1}{p}} D_p(\cdot)$, define

$$J : (0, \infty) \times (0, \infty) \to \mathbb{R}, \qquad (x_1, x_2) \mapsto \frac{x_1^{1/p}}{x_2},$$

(5.8)

$$f_1(x) = e^{px}, \qquad f_2(x) = e^x.$$

(c) For the normalized geometric mean process $nS(\cdot)$, pick

$$J : \mathbb{R} \times (0, \infty) \to \mathbb{R}, \qquad (x_1, x_2) \mapsto \frac{e^{x_1}}{x_2},$$

(5.9)

$$f_1(x) = x, \qquad f_2(x) = e^x.$$

It is elementary to check the assumptions of Theorem 1.4 for all of these functions. □

5.2. *Hitting times of the market diversity.* In this subsection, we specialize Theorem 1.5 to the measures of diversity from Section 5.1. To this end, we denote by $H^*(\cdot)$, $D_p^*(\cdot)$, $p \in (0, 1)$ and $S^*(\cdot)$ the limiting entropy, ℓ^p-norms and geometric mean processes, respectively,

$$H^*(\cdot) = \log\langle e^x, \varrho(\cdot)\rangle - \frac{\langle x e^x, \varrho(\cdot)\rangle}{\langle e^x, \varrho(\cdot)\rangle},$$

(5.10)
$$D_p^*(\cdot) = \frac{\langle e^{px}, \varrho(\cdot)\rangle^{1/p}}{\langle e^x, \varrho(\cdot)\rangle}, \qquad p \in (0, 1),$$

$$S^*(t) = \frac{e^{\langle x, \varrho(\cdot)\rangle}}{\langle e^x, \varrho(\cdot)\rangle}.$$

The functions J, $f_1, \ldots, f_k$ in the three cases, given explicitly in (5.7), (5.8) and (5.9), respectively, satisfy the assumptions in Theorem 1.5, so that one only needs to verify the statements in (1.12) for the coefficients b, σ and levels a of interest. The next proposition provides the dynamics of $H^*(\cdot)$, $D_p^*(\cdot)$, $p \in (0, 1)$ and $S^*(\cdot)$, thus, yielding a sufficient condition on b, σ and a for Theorem 1.5 to apply.

PROPOSITION 5.4. *Under Assumption* 1.1, *consider*

$$(5.11) \qquad \varrho_p(\cdot) := \frac{e^{px} \varrho(\cdot)}{\langle e^{px}, \varrho(\cdot) \rangle} \in C\big([0, \infty), M_1(\mathbb{R})\big), \quad p \in [0, 1).$$

Then one has for the processes $H^*(\cdot)$, $D_p^*(\cdot)$, $p \in (0, 1)$ and $S^*(\cdot)$ of (5.10):

$$(5.12) \qquad \begin{aligned} \frac{\mathrm{d}H^*(t)}{\mathrm{d}t} &= -\frac{1}{2}\big\langle \sigma\big(R(t, \cdot)\big)^2, \varrho_1(t) \big\rangle \\ &\quad - \mathrm{cov}_{\varrho_1(t)}\bigg(x, b\big(R(t, \cdot)\big) + \frac{\sigma(R(t, \cdot))^2}{2} \bigg), \end{aligned}$$

$$(5.13) \qquad \begin{aligned} \frac{\mathrm{d}D_p^*(t)}{\mathrm{d}t} &= D_p^*(t)\bigg(\bigg\langle b\big(R(t, \cdot)\big) + \frac{p\sigma(R(t, \cdot))^2}{2}, \varrho_p(t) \bigg\rangle \\ &\quad - \bigg\langle b\big(R(t, \cdot)\big) + \frac{\sigma(R(t, \cdot))^2}{2}, \varrho_1(t) \bigg\rangle \bigg), \end{aligned}$$

$$(5.14) \qquad \begin{aligned} \frac{\mathrm{d}S^*(t)}{\mathrm{d}t} &= S^*(t)\bigg(\big\langle b\big(R(t, \cdot)\big), \varrho(t) \big\rangle \\ &\quad - \bigg\langle b\big(R(t, \cdot)\big) + \frac{\sigma(R(t, \cdot))^2}{2}, \varrho_1(t) \bigg\rangle \bigg). \end{aligned}$$

In particular, whenever $b + \frac{\sigma^2}{2}$ is an increasing function and $a \in (-\infty, H^*(0)]$, $a \in (0, D_p^*(0)]$, $p \in (0, 1)$ or $a \in (0, S^*(0)]$, the assertions in (1.12) *hold, and hence, also the conclusion of Theorem* 1.5 *for the resulting hitting times of the normalized processes $H(\cdot) - \log n$, $n^{\frac{p-1}{p}} D_p(\cdot)$, $p \in (0, 1)$ or $n S(\cdot)$.*

REMARK 5.5. More specifically, whenever $b + \frac{\sigma^2}{2}$ is an increasing function and $a \in (-\infty, H^*(0)]$, $a \in (0, D_p^*(0)]$, $p \in (0, 1)$ or $a \in (0, S^*(0)]$, the conclusion of Theorem 1.5 reads as follows (recall (5.7), (5.8), (5.9)):

$$\begin{aligned} &\frac{\mathrm{d}H^*(\tau_H)}{\mathrm{d}t} \cdot \sqrt{n}\big(\tau_H^{(n)} - \tau_H\big) \\ &\overset{n \to \infty}{\longrightarrow} \bigg(\frac{1}{\langle e^x, \varrho(\tau_H) \rangle} + \frac{\langle x e^x, \varrho(\tau_H) \rangle}{\langle e^x, \varrho(\tau_H) \rangle^2} \bigg) \int_{\mathbb{R}} e^x G(\tau_H, x)\, \mathrm{d}x \\ &\quad - \frac{1}{\langle e^x, \varrho(\tau_H) \rangle} \int_{\mathbb{R}} e^x (x + 1) G(\tau_H, x)\, \mathrm{d}x, \end{aligned}$$

$$\frac{\mathrm{d}D_p^*(\tau_D)}{\mathrm{d}t} \cdot \sqrt{n}\big(\tau_D^{(n)} - \tau_D\big)$$

$$\overset{n\to\infty}{\longrightarrow} \frac{\langle e^{px}, \varrho(\tau_D)\rangle^{1/p-1}}{\langle e^x, \varrho(\tau_D)\rangle} \int_{\mathbb{R}} e^{px} G(\tau_D, x)\, \mathrm{d}x$$

$$- \frac{\langle e^{px}, \varrho(\tau_D)\rangle^{1/p}}{\langle e^x, \varrho(\tau_D)\rangle^2} \int_{\mathbb{R}} e^x G(\tau_D, x)\, \mathrm{d}x,$$

$$\frac{\mathrm{d}S^*(\tau_S)}{\mathrm{d}t} \cdot \sqrt{n}\big(\tau_S^{(n)} - \tau_S\big)$$

$$\overset{n\to\infty}{\longrightarrow} \frac{e^{\langle x, \varrho(\tau_S)\rangle}}{\langle e^x, \varrho(\tau_S)\rangle} \int_{\mathbb{R}} G(\tau_S, x)\, \mathrm{d}x$$

$$- \frac{e^{\langle x, \varrho(\tau_S)\rangle}}{\langle e^x, \varrho(\tau_S)\rangle^2} \int_{\mathbb{R}} e^x G(\tau_S, x)\, \mathrm{d}x,$$

where $\frac{\mathrm{d}H^*(\tau_H)}{\mathrm{d}t}$, $\frac{\mathrm{d}D_p^*(\tau_H)}{\mathrm{d}t}$, $\frac{\mathrm{d}S^*(\tau_H)}{\mathrm{d}t}$ are given by (5.13), (5.14), (5.15), while $\tau_H^{(n)}$, τ_H, $\tau_D^{(n)}$, τ_D, $\tau_S^{(n)}$, τ_S are the hitting times of $H(\cdot) - \log n$, $H^*(\cdot)$, $n^{\frac{p-1}{p}} D_p(\cdot)$, $D_p^*(\cdot)$, $nS(\cdot)$, $S^*(\cdot)$, respectively.

REMARK 5.6. It is worth mentioning that the process of push-forwards of $\varrho_p(t)$, $t \geq 0$ under $R(t, \cdot)$, $t \geq 0$, respectively, has appeared previously in [19] under the name "asymptotic weighted capital measure." Therein, it plays a crucial role in the analysis of the long-term behavior of the capital distribution in large rank-based models. Remarkably, a monotonicity assumption on $b + \frac{\sigma^2}{2}$, as in Proposition 5.4, also appears in [19]. In [19], Corollary 1, it is shown to imply the monotonicity in p of the limiting (as $n \to \infty$) long-term growth rate of the portfolio generated by $D_p(\cdot)$. As our proof of Proposition 5.4 below reveals, the assumption of an increasing $b + \frac{\sigma^2}{2}$ ensures that the limiting market diversity, as measured by $H^*(\cdot)$, $D_p^*(\cdot)$, $p \in (0, 1)$ or $S^*(\cdot)$, is decreasing in time, sometimes a desirable modeling assumption.

We prepare the following continuous version of Chebyshev's sum inequality for the proof of Proposition 5.4.

LEMMA 5.7. *For all $\nu \in M_1(\mathbb{R})$ and increasing functions f, g on $\mathbb{R}$ integrable with respect to ν,*

(5.15) $$\langle fg, \nu\rangle \geq \langle f, \nu\rangle \langle g, \nu\rangle.$$

PROOF. Since g is increasing, there exists an $x_0 \in \mathbb{R}$ such that $g(x) \leq \langle g, \nu\rangle$ if $x < x_0$ and $g(x) \geq \langle g, \nu\rangle$ if $x > x_0$. By distinguishing between $x < x_0$ and $x \geq x_0$

and using that f is increasing, we deduce

$$(5.16) \qquad f(x)\big(g(x) - \langle g, \nu \rangle\big) \geq f(x_0)\big(g(x) - \langle g, \nu \rangle\big), \quad x \in \mathbb{R}.$$

Integrating both sides with respect to ν and rearranging we arrive at (5.15). $\qquad \square$

We are now ready to present the proof of Proposition 5.4.

PROOF OF PROPOSITION 5.4.   We recall the notation $\mathcal{A}_t = b(R(t, \cdot))\frac{\mathrm{d}}{\mathrm{d}x} + \frac{\sigma(R(t,\cdot))^2}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}$, $t \geq 0$ and that, for any $f \in \mathcal{E}_2$,

$$(5.17) \qquad \langle f, \varrho(t) \rangle - \langle f, \varrho(0) \rangle = \int_0^t \langle \mathcal{A}_s f, \varrho(s) \rangle \, \mathrm{d}s, \quad t \geq 0$$

(cf. (4.34)). Due to the boundedness of b, σ (cf. Assumption 1.1(b)) and $f \in \mathcal{E}_2$ the dominated convergence theorem implies that the function $s \mapsto \langle \mathcal{A}_s f, \varrho(s) \rangle$ is continuous on $[0, \infty)$, and thus,

$$(5.18) \qquad \frac{\mathrm{d}\langle f, \varrho(t) \rangle}{\mathrm{d}t} = \langle \mathcal{A}_t f, \varrho(t) \rangle, \quad t \geq 0.$$

Therefore, in the setting of Theorem 1.5,

$$(5.19) \qquad \frac{\mathrm{d}\mathcal{J}_{J; f_1, \dots, f_k}(\varrho(t))}{\mathrm{d}t} = \sum_{j=1}^k \mathcal{J}_{J_{x_j}; f_1, \dots, f_k}(\varrho(t))\langle \mathcal{A}_t f_j, \varrho(t) \rangle, \quad t \geq 0.$$

To obtain the differential equations (5.13), (5.14) and (5.15) it suffices to insert into (5.19) the formulas from (5.7), (5.8) and (5.9), respectively, and to simplify the result.

Supposing, in addition, that $b + \frac{\sigma^2}{2}$ is increasing we can first employ Lemma 5.7 with $\nu = \varrho_1(t)$, $f(x) = x$ and $g(x) = b(R(t, x)) + \frac{\sigma(R(t,x))^2}{2}$ to find

$$(5.20) \qquad \mathrm{cov}_{\varrho_1(t)}\left(x, b\big(R(t, \cdot)\big) + \frac{\sigma(R(t, \cdot))^2}{2}\right) \geq 0, \quad t \geq 0.$$

Consequently, we read off from (5.13) that

$$(5.21) \qquad \frac{\mathrm{d}H^*(t)}{\mathrm{d}t} \leq -\frac{1}{2} \min_{r \in [0,1]} \sigma(r)^2, \quad t \geq 0,$$

so (1.12) must hold for all $a \in (-\infty, H^*(0)]$, and the conclusion of Theorem 1.5 applies to the hitting times of such a by $H(\cdot) - \log n$.

Now, we take $\nu = \varrho_p(t)$, $f(x) = b(R(t, x)) + \frac{\sigma(R(t,x))^2}{2}$ and $g(x) = e^{(1-p)x}$, with $p \in [0, 1)$, in Lemma 5.7 to get

$$
(5.22) \quad
\begin{aligned}
&\left\langle b\big(R(t, \cdot)\big) + \frac{\sigma(R(t, \cdot))^2}{2}, \varrho_1(t) \right\rangle \\
&\qquad \geq \left\langle b\big(R(t, \cdot)\big) + \frac{\sigma(R(t, \cdot))^2}{2}, \varrho_p(t) \right\rangle, \quad t \geq 0.
\end{aligned}
$$

The values of $p \in (0, 1)$ and $p = 0$ reveal

(5.23)
$$\frac{d \log D_p^*(\cdot)}{dt} \leq -\frac{1-p}{2} \min_{r \in [0,1]} \sigma(r)^2, \quad p \in (0, 1) \quad \text{and}$$
$$\frac{d \log S^*(\cdot)}{dt} \leq -\frac{1}{2} \min_{r \in [0,1]} \sigma(r)^2,$$

respectively, yielding the remaining assertions. □

REMARK 5.8.   A verification of the conditions in (1.12) beyond the setup in Proposition 5.4 seems to require information on $\varrho(\cdot)$ or, equivalently, R that needs to be deduced on a case-by-case basis. This is possible, for example, when (1.3) is of the special form

(5.24) $$dX_i^{(n)}(t) = \big(2C_1 F_{\varrho^{(n)}(t)}(X_i^{(n)}(t)) + C_2\big) dt + \sigma\, dB_i^{(n)}(t), \quad i = 1, 2, \ldots, n$$

for some $C_1 \neq 0$, $C_2 \in \mathbb{R}$ and $\sigma > 0$. Indeed, then the Cauchy problem for the porous medium equation (1.5) reduces to the one for the generalized Burgers equation

(5.25) $$R_t = -(C_1 R^2 + C_2 R)_x + \frac{\sigma^2}{2} R_{xx}, \qquad R(0, \cdot) = F_\lambda(\cdot).$$

The solution of the latter is provided by the Cole–Hopf transformation $R = -\frac{2C_1}{\sigma^2}(\log \varphi)_x$, where φ is the solution of the Cauchy problem for the heat equation

(5.26) $$\varphi_t = -C_2 \varphi_x + \frac{\sigma^2}{2} \varphi_{xx}, \qquad \varphi(0, x) = e^{-\frac{\sigma^2}{2C_1} \int_0^x F_\lambda(y)\, dy}.$$

For any fixed $\lambda \in M_1(\mathbb{R})$ (perhaps retrieved from the observed market capitalizations), φ is given explicitly by a convolution with the heat kernel, and one can check if the conditions in (1.12) are valid for the resulting $\varrho(\cdot) = R_x(\cdot, x)\, dx$.

5.3. *Performance of functionally generated portfolios.*   This last subsection is devoted to a discussion of the performance of multiplicatively and additively generated portfolios $\pi^{\widetilde{\Psi}; \times}$ and $\pi^{\widetilde{\Psi}; +}$, as defined in the Introduction. We focus initially on their associated nondecreasing *excess growth* processes

(5.27)
$$-\frac{1}{2} \sum_{i,j=1}^n \int_0^t \frac{\widetilde{\Psi}_{x_i x_j}(\mu(\cdot))}{\widetilde{\Psi}(\mu(\cdot))}\, d[\mu_i, \mu_j](\cdot), \quad t \geq 0 \quad \text{and}$$
$$-\frac{1}{2} \sum_{i,j=1}^n \int_0^t \widetilde{\Psi}_{x_i x_j}(\mu(\cdot))\, d[\mu_i, \mu_j](\cdot), \quad t \geq 0$$

which enter the value processes $V^{\widetilde{\Psi}; \times}(\cdot)$ and $V^{\widetilde{\Psi}; +}(\cdot)$ relative to that of the market portfolio $\mu(\cdot)$ according to (1.16) and (1.17), respectively. Under the assumptions

in (1.21), as well as (1.22) or (1.26), respectively, the former obey the concentration of measure estimate from [16], Corollary 8 (note the symmetry of $\widetilde{\Psi}$ due to (1.20), (1.18); the strong law of large numbers for the process $\mu(\cdot)$ in [2], equation (4.5); and that the derivation of [16], Corollary 8, for multiplicatively generated portfolios carries over mutatis mutandis to the case of additive generation).

PROPOSITION 5.9. *Suppose the assumptions in* (1.21), *as well as* (1.22) *or* (1.26) *are satisfied for some* $n \in \mathbb{N}$. *Then, for all* $r, t, \varepsilon > 0$ *and in the notation of Corollary* 1.8, *Remark* 1.10,

$$(5.28) \quad \begin{aligned} &\mathbb{P}\left(-\frac{1}{2t}\sum_{i,j=1}^{n}\int_{0}^{t}\frac{\widetilde{\Psi}_{x_i x_j}(\mu(\cdot))}{\widetilde{\Psi}(\mu(\cdot))}\,\mathrm{d}[\mu_i, \mu_j](\cdot) \leq r^{\times} - r\right) \\ &\leq \left\|\frac{\mathrm{d}\kappa^{(n)}}{\mathrm{d}\zeta^{(n)}}\right\|_{L^2(\zeta^{(n)})}e^{-c^{\times}(r,\varepsilon)t} \end{aligned}$$

or

$$(5.29) \quad \begin{aligned} &\mathbb{P}\left(-\frac{1}{2t}\sum_{i,j=1}^{n}\int_{0}^{t}\widetilde{\Psi}_{x_i x_j}(\mu(\cdot))\,\mathrm{d}[\mu_i, \mu_j](\cdot) \leq r^{+} - r\right) \\ &\leq \left\|\frac{\mathrm{d}\kappa^{(n)}}{\mathrm{d}\zeta^{(n)}}\right\|_{L^2(\zeta^{(n)})}e^{-c^{+}(r,\varepsilon)t}, \end{aligned}$$

respectively. Hereby,

$$\begin{aligned} &c^{+}(r, \varepsilon) \\ &:= \frac{\min_{1\leq j\leq n-1}(\sum_{i=1}^{j}b(\frac{i}{n}) - \frac{j}{n}\sum_{i=1}^{n}b(\frac{i}{n}))^2}{2 - 2\cos\frac{\pi}{n}} \\ &\quad \cdot \max\left(\frac{r^2}{(C^+)^2}, \\ &\qquad 4\varepsilon(\varepsilon + v^+)\left(\sqrt{1 + \frac{r^2}{2\varepsilon(\varepsilon + v^+)^2\max(|C^{+,\uparrow}|, |C^{+,\downarrow}|)^2}} - 1\right)\right), \end{aligned}$$

with C^+, $C^{+,\uparrow}$, $C^{+,\downarrow}$ *and* v^+ *being the overall expression, the essential supremum, the essential infimum and the variance under* $\zeta^{(n)}$ *of the expression inside the essential supremum in* (1.26).

If, in addition, Assumption 1.1 holds, one can combine Proposition 5.9 with Theorem 1.4 by using the union bound and obtain, for all $r, s, t, \varepsilon > 0$, the perfor-

mance estimates

$$(5.30) \qquad \mathbb{P}\left(V^{\widetilde{\Psi};\times}(t) \le \frac{\mathcal{J}_{J;f_1,\dots,f_k}(\varrho(t)) - s/\sqrt{n}}{\widetilde{\Psi}(\mu(0))} e^{(r^\times - r)t}\right)$$

$$\le \overline{\Phi}(s/\chi_t)(1 + o_n(1)) + \left\|\frac{d\kappa^{(n)}}{d\zeta^{(n)}}\right\|_{L^2(\zeta^{(n)})} e^{-c^\times(r,\varepsilon)t}$$

or

$$(5.31) \qquad \mathbb{P}(V^{\widetilde{\Psi};+}(t) \le 1 + \mathcal{J}_{J;f_1,\dots,f_k}(\varrho(t)) - s/\sqrt{n} - \widetilde{\Psi}(\mu(0)) + (r^+ - r)t)$$

$$\le \overline{\Phi}(s/\chi_t)(1 + o_n(1)) + \left\|\frac{d\kappa^{(n)}}{d\zeta^{(n)}}\right\|_{L^2(\zeta^{(n)})} e^{-c^+(r,\varepsilon)t},$$

respectively, where $\overline{\Phi}$ is the standard normal tail cumulative distribution function, χ_t is the standard deviation of the time t value of the Gaussian process on the right-hand side of (1.11) and $o_n(1)$ is a quantity tending to 0 as $n \to \infty$. Complementary to the performance estimates (5.30), (5.31) for *fixed* times, Corollary 1.8, which is proved next, provides a bound on the *random* time it takes to reach the desired performance.

PROOF OF COROLLARY 1.8. Our starting point is the observation that

$$(5.32) \qquad \mathbb{P}(\eta^{\widetilde{\Psi};\times} \ge \tau + s/\sqrt{n}) \le \mathbb{P}(\eta^{\widetilde{\Psi};\times} \ge \tau^{(n)}) + \mathbb{P}(\tau^{(n)} \ge \tau + s/\sqrt{n}).$$

By the definition of $\eta^{\widetilde{\Psi};\times}$, the first of the latter two summands is less or equal to

$$(5.33) \qquad \mathbb{P}\left(V^{\widetilde{\Psi};\times}(\tau^{(n)}) \le \frac{a}{\widetilde{\Psi}(\mu(0))} e^{(r^\times - r)(\tau - s/\sqrt{n})}\right)$$

$$= \mathbb{P}\left(-\frac{1}{2}\sum_{i,j=1}^n \int_0^{\tau^{(n)}} \frac{\widetilde{\Psi}_{x_i x_j}(\mu(\cdot))}{\widetilde{\Psi}(\mu(\cdot))} d[\mu_i, \mu_j](\cdot) \le (r^\times - r)(\tau - s/\sqrt{n})\right)$$

(recall (1.16) and $\widetilde{\Psi}(\mu(\tau^{(n)})) = a$). Since the excess growth process is nondecreasing, the probability on the right-hand side of (5.33) is at most

$$(5.34) \qquad \mathbb{P}(\tau^{(n)} \le \tau - s/\sqrt{n})$$

$$+ \mathbb{P}\left(-\frac{1}{2}\sum_{i,j=1}^n \int_0^{\tau - s/\sqrt{n}} \frac{\widetilde{\Psi}_{x_i x_j}(\mu(\cdot))}{\widetilde{\Psi}(\mu(\cdot))} d[\mu_i, \mu_j](\cdot) \le (r^\times - r)(\tau - s/\sqrt{n})\right).$$

Using (1.14) for the second summand on the right-hand side of (5.32) and the first summand in (5.34), then (5.28) for the second summand in (5.34) we get (1.24). □

REFERENCES

[1] ALDOUS, D. J. (2002). How many Brownian particles escape when you control with total drift = 1? Available at stat.berkeley.edu/~aldous/Research/OP/river.pdf.

[2] BANNER, A. D., FERNHOLZ, R. and KARATZAS, I. (2005). Atlas models of equity markets. *Ann. Appl. Probab.* **15** 2296–2330. MR2187296

[3] BASS, R. F. and PARDOUX, É. (1987). Uniqueness for diffusions with piecewise constant coefficients. *Probab. Theory Related Fields* **76** 557–572. MR0917679

[4] BOBKOV, S. and LEDOUX, M. (2016). One-dimensional empirical measures, order statistics, and kantorovich transport distances. Available at perso.math.univ-toulouse.fr/ledoux/files/2016/12/MEMO.pdf.

[5] BOGACHEV, V. I. (2007). *Measure Theory. Vol. I, II*. Springer, Berlin. MR2267655

[6] CABEZAS, M., DEMBO, A., SARANTSEV, A. and SIDORAVICIUS, V. (2017). Brownian particles with rank-dependent drifts: Out-of-equilibrium behavior. Available at arXiv.org/abs/1708.01918.

[7] DEMBO, A., JARA, M. and OLLA, S. (2017). The infinite atlas process: Convergence to equilibrium. Available at arXiv.org/abs/1709.04085.

[8] DEMBO, A., SHKOLNIKOV, M., VARADHAN, S. R. S. and ZEITOUNI, O. (2016). Large deviations for diffusions interacting through their ranks. *Comm. Pure Appl. Math.* **69** 1259–1313. MR3503022

[9] DEMBO, A. and TSAI, L.-C. (2017). Equilibrium fluctuation of the Atlas model. *Ann. Probab.* **45** 4529–4560. MR3737917

[10] FERNHOLZ, E. R. (2002). *Stochastic Portfolio Theory: Stochastic Modelling and Applied Probability. Applications of Mathematics* (*New York*) **48**. Springer, New York. MR1894767

[11] FERNHOLZ, .R. and KARATZAS, I. (2009). *Stochastic Portfolio Theory: An Overview. Handbook of Numerical Analysis. Mathematical Modeling and Numerical Methods in Finance*.

[12] FERNHOLZ, R. (1999). On the diversity of equity markets. *J. Math. Econom.* **31** 393–417. MR1684221

[13] FERNHOLZ, R., KARATZAS, I. and KARDARAS, C. (2005). Diversity and relative arbitrage in equity markets. *Finance Stoch.* **9** 1–27. MR2210925

[14] GILDING, B. H. (1989). Improved theory for a nonlinear degenerate parabolic equation. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (4) **16** 165–224. MR1041895

[15] HAJEK, B. (1985). Mean stochastic comparison of diffusions. *Z. Wahrsch. Verw. Gebiete* **68** 315–329. MR0771469

[16] ICHIBA, T., PAL, S. and SHKOLNIKOV, M. (2013). Convergence rates for rank-based models with applications to portfolio theory. *Probab. Theory Related Fields* **156** 415–448. MR3055264

[17] ICHIBA, T., PAPATHANAKOS, V., BANNER, A., KARATZAS, I. and FERNHOLZ, R. (2011). Hybrid atlas models. *Ann. Appl. Probab.* **21** 609–644. MR2807968

[18] JOURDAIN, B. and REYGNER, J. (2013). Propogation of chaos for rank-based interacting diffusions and long time behaviour of a scalar quasilinear parabolic equation. *Stoch. Partial Differ. Equ. Anal. Comput.* **1** 455–506. MR3327514

[19] JOURDAIN, B. and REYGNER, J. (2015). Capital distribution and portfolio performance in the mean-field Atlas model. *Ann. Finance* **11** 151–198. MR3340241

[20] KARATZAS, I. and RUF, J. (2017). Trading strategies generated by Lyapunov functions. *Finance Stoch.* **21** 753–787. MR3663643

[21] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. Springer, New York. MR1121940

[22] KOLLI, P. and SHKOLNIKOV, M. (2018). SPDE limit of the global fluctuations in rank-based models. *Ann. Probab.* **46** 1042–1069. MR3773380

[23] KRYLOV, N. V. (1971). A certain estimate from the theory of stochastic integrals. *Teor. Veroyatn. Primen.* **16** 446–457. MR0298792

[24] MCKEAN, H. P. and SHEPP, L. A. (2005). The advantage of capitalism vs. socialism depends on the criterion. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov.* (*POMI*) **328** 160–168, 279–280. MR2214539

[25] PAL, S. and PITMAN, J. (2008). One-dimensional Brownian particle systems with rank-dependent drifts. *Ann. Appl. Probab.* **18** 2179–2207. MR2473654

[26] REYGNER, J. (2015). Chaoticity of the stationary distribution of rank-based interacting diffusions. *Electron. Commun. Probab.* **20** no. 60, 20. MR3399811

[27] SARANTSEV, A. and TSAI, L.-C. (2017). Stationary gap distributions for infinite systems of competing Brownian particles. *Electron. J. Probab.* **22** Paper No. 56, 20. MR3672832

[28] SHKOLNIKOV, M. (2012). Large systems of diffusions interacting through their ranks. *Stochastic Process. Appl.* **122** 1730–1747. MR2914770

[29] STROOCK, D. W. and VARADHAN, S. R. S. (2006). *Multidimensional Diffusion Processes*. *Classics in Mathematics*. Springer, Berlin. MR2190038

[30] SZNITMAN, A.-S. (1991). Topics in propagation of chaos. In *École D'Été de Probabilités de Saint-Flour XIX—1989*. *Lecture Notes in Math.* **1464** 165–251. Springer, Berlin. MR1108185

[31] TANG, W. and TSAI, L.-C. (2018). Optimal surviving strategy for drifted Brownian motions with absorption. *Ann. Probab.* **46** 1597–1650. MR3785596

S. A. ALMADA MONTER
QUANTITATIVE RESEARCH
JPMORGAN CHASE
NEW YORK, NEW YORK 10017
USA
E-MAIL: saalm56@gmail.com

M. SHKOLNIKOV
ORFE DEPARTMENT
BENDHEIM CENTER FOR FINANCE,
   AND PROGRAM IN APPLIED &
   COMPUTATIONAL MATHEMATICS
PRINCETON UNIVERSITY
PRINCETON, NEW JERSEY 08544
USA
E-MAIL: mshkolni@gmail.com

J. ZHANG
ORFE DEPARTMENT
PRINCETON UNIVERSITY
PRINCETON, NEW JERSEY 08544
USA
E-MAIL: jiacheng@princeton.edu